The von Neumann Minimax Theorem and Its Relatives and A Study of Externality in On-line Auctions

Samaneh Malekan

Mathematics

A thesis

Submitted in partial fulfillment

of the requirements for the degree of

Master of Science

Faculty of Mathematics and Science, Brock University

St. Catharines, Ontario

Abstract

This work consists of a theoretical part and an experimental one. The first part provides a simple treatment of the celebrated von Neumann minimax theorem as formulated by Nikaidô and Sion. It also discusses its relationships with fundamental theorems of convex analysis.

The second part is about externality in sponsored search auctions. It shows that in these auctions, advertisers have externality effects on each other which influence their bidding behavior. It proposes Hal R.Varian model and shows how adding externality to this model will affect its properties. In order to have a better understanding of the interaction among advertisers in on-line auctions, it studies the structure of the Google advertisements network and shows that it is a small-world scale-free network.

Acknowledgements

I would like to extend my sincere gratitude and appreciation to all those who made this M.Sc. thesis possible. Special thanks are due to my supervisors, Dr. Hichem Ben-El-Mechaiekh and Dr. Babak Farzad, for their support and guidance during this research. Special thanks to the Department of Mathematics at the Brock University.

I am deeply indebted to my family and friends for their supports. I would like to thank my family for their help and encouragement. My mom, and my dad, have inspired me with their achievements, and instilled in me with confidence to set the bar high, and go for it. My lovely sister and her awesome husband, it was a great feeling to know there are two great persons, not so far away, that will accept me with open arms and all the love.

Samaneh Malekan

September 2010

Contents

Chapter 1

Introduction

The first part of this work, which is theoretical in nature, is covered in Chapter 2 and presents an elementary treatment of the celebrated von Neumann minimax theorem for quasiconcave/convex and lower/upper semicontinuous functions as formulated by Nikaidô and Sion [35, 39]. It also discusses its relationships with fixed point and coincidence theorems for set-valued maps (see e.g., [6, 9, 15, 16]), the Knaster-Kuratowski-Mazurkiewicz principle (KKM) principle [29], and system of nonlinear inequalities. All are fundamental results in convex analysis.

The motivation of our approach is to provide a simple and direct proof of the Nikaidô and Sion formulation of the minimax theorem that is readily accessible to readers without extensive background in functional analysis and topology and to shed some light on its relationships with landmark results of convex analysis. Following Ben-El-Mechaiekh and Dimand [10], we base our argument for the elementary proof on a result of Victor Klee [28] and Claude Berge [13] on convex covers of closed convex subsets of Euclidean spaces. The proof of the Berge-Klee result is itself based on the simplest version of the theorem on the separation of convex sets in Euclidean spaces, which is accessible to students in a first course on continuous optimization.

The literature on the von Neumann minimax theorem is quite extensive and includes a number of different proofs. Those based on the KKM lemma or the Brouwer

fixed point theorem are short (see e.g., [9, 12, 15, 16, 27]) but they need preparatory work beyond the courses of a typical undergraduate program in mathematics (e.g. such as Sperner's lemma on the existence of a complete labeling for a Euclidean simplex, or the non retraction theorem of the unit ball onto its boundary in a Euclidean space, or homology and homotopy theories, ...). Also, a number of elementary proofs were provided for the minimax theorem. It is worth mentioning the proof given by I. Joó [25] based on a result of F. Riesz on the nonempty intersection of a family of compact sets having the finite intersection property. Another elementary but not simple proof is due to J. Kindler and is based on the 1-dimensional KKM theorem, the 1-dimensional Helly theorem (i.e., any family of pairwise intersecting compact intervals in R has nonempty intersection), and Zorn's lemma [26].
We truly believe that the proof presented here is the simplest in the literature.

We were also interested in the relationships between fundamental results in convex functional analysis as illustrated by the chart in Figure 1.1, with a particular attention to the problem: does the Berge-Klee intersection theorem (or any Helly type intersection theorem) imply the KKM principle or one of its relatives (Sperner lemma, Matching Theorem of Ky Fan, etc.)?

One can see that many of the results in this chart are equivalent. We shall include the proofs of the solid arrows but not those of dashed ones, simply referring to the articles where they have appeared. The implications 1-4 and the circular tour in the lower loop of the flow-chart are much in the spirit of [11, 10] . Implication 11 extends the "elementary KKM principle" of [23] to arbitrary topological vector spaces.

It is worth mentioning that it is still an open question as to whether or not the Berge-Klee intersection theorem implies the KKM principle (or even its ancestor, the finite dimensional Sperner's lemma). A positive answer would yield a truly elementary proof of the KKM principle. It is however known [24] that a "topological Klee" is equivalent to the KKM Principle. But the proof of the topological Klee involves the KKM principle. In other words, the jury is still out on an elementary proof of the

KKM principle. The bold arrow 13 in Figure 1.1 is thus still an open question.

The tour proposed here affords the reader the opportunity to start a discussion on the minimax theorem from any of the cells of the diagram.

In *game theory*, a *two player zero-sum game* is a game in which a participant's gain or loss is exactly balanced by the losses or gains of the other participant. In other words, the sum of the gain of one participant and the loss of the other one is zero. A two player zero-sum game can be viewed mathematically as triple (X, Y, f) where X is the set of strategies of player 1, Y is the set of strategies of player 2, and $f : X \times Y \to \mathbb{R}$ is a mapping, called the *pay-off function*. For each $x \in X$ and each $y \in Y$, $f(x, y)$ is what player 1 gains and what player 2 looses when they choose strategies x and y, respectively. Both players know the payoffs associated to secretly chosen strategies, but each player is unaware of the other player's choice. Then, the choices are revealed and each player's points total is affected according to the payoffs for those choices.

Consider two player zero-sum games with arbitrary numbers of finite strategies. Each player aims at choosing so-called *security strategies*, which for player 1 are of *maxmin* type: the number $\alpha := \max_X \min_Y f(x, y)$ represents his guaranteed minimal gain; while for player 2, they are of *minmax* type: the number $\beta := \min_Y \max_X f(x, y)$ being his guaranteed maximal loss. It is easily seen that always, $\alpha \leq \beta$. When equality holds, the common value $\alpha = \beta$ is known as the *value of the game*.

Emile Borel was the first to raise the problem as to whether or not such a zero-sum games always have a value (see [12]). In 1928, John von Neumann theorem answered Borel's question in the affirmative (without providing the proof) for arbitrary finite strategy sets [41] with what was to become the celebrated von Neumann minimax theorem. This result is the basis of game theory as a distinct mathematical discipline. The theorem soon was generalized to games with n players, non-constant sums of payoffs, infinitely many players, etc ... (see [12, 13, 18] for detailed historical accounts).

Von Neumann published the first proof in *Mathematical Annalen* later in 1928 [43]. He proposed in 1932 a second proof based on Brouwer's fixed point theorem at the Menger's Colloquium in Vienna [42]. Although von Neumann's originally formulated the minimax theorem for linear forms, he quickly became well aware that convexity of level sets of the functionals involved was sufficient for the proof to hold true. It was not until 1954 that H. Nikaidô [35] and later in 1958 M. Sion [39] formulated the minimax theorem for quasiconvex/concave and lower/upper semicontinuous functions.

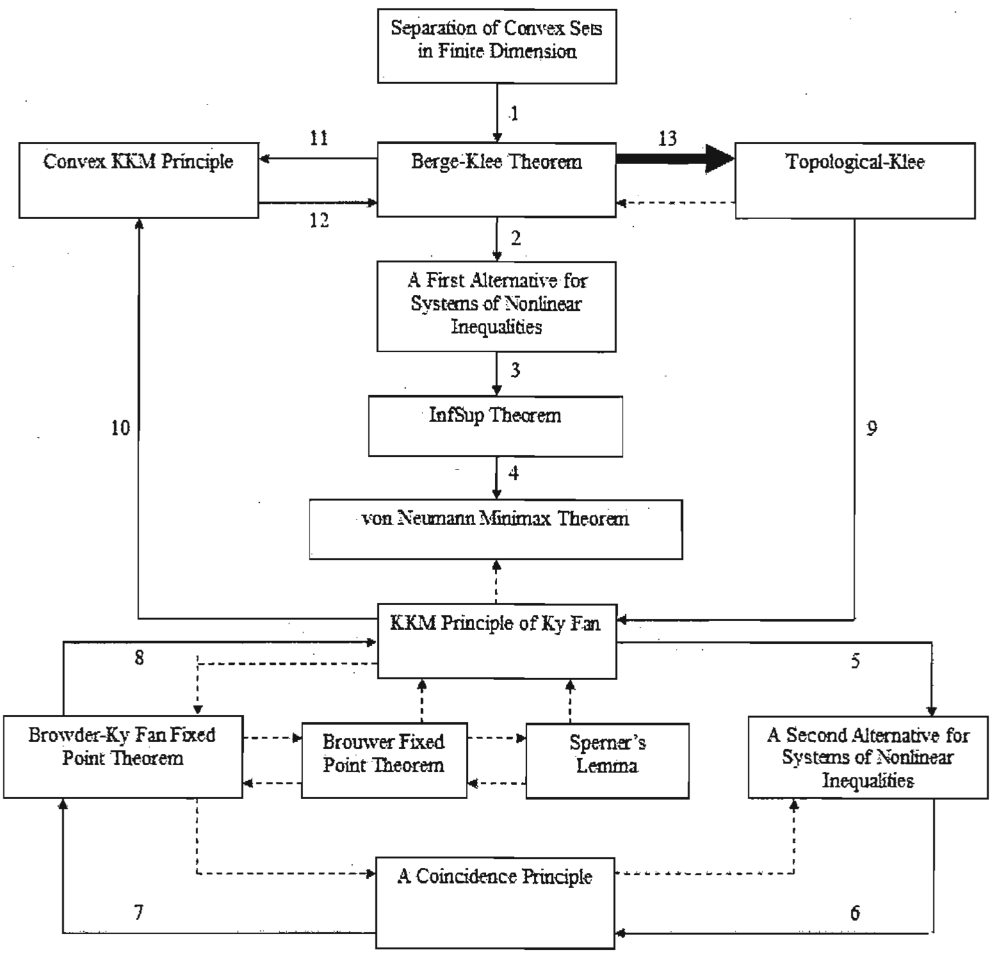

Figure 1.1: relationship between fundamental results in convex analysis.

In the first section of chapter 2, we propose some definitions, examples, and preliminary results culminating with the theorem on the separation of disjoint closed convex subsets of Euclidean spaces together with a most elementary proof. The core of this chapter is presented in section 2 where we prove the various implications in Figure 1.1, including the simple and elementary proof, due to Ben-El-Mechaiekh and Dimand [10], of Nikaidô-Sion formulation of the minimax theorem.

The second part of the thesis, described in Chapters 3 and 4, is experimental. The existence of Nash equilibria for games, which is known to be equivalent to the von Neumann minimax theorem, is at the heart of the solvability of online auction models studied in this second part. Chapter 3 is about sponsored search, a service provided by search engines. Sponsored search is the delivery of web addresses with advertisements relevant to the user's inquiry as part of the search experience. Search engines, such as Google, Yahoo! and MSN use sponsored search to earn money by displaying advertisements alongside the result page of a user inquiry.

Sponsored search satisfies users' desire for relevant search results and advertisers need for making traffic to their web sites. It is now considered to be among the most effective marketing tools available. Sponsored search has become a big business among different kinds of advertising methods. For example, Google generated roughly $22 billion in revenue in 2009 which is almost 97% of its revenue.

Search engines allocate certain spaces in the result page of an inquiry to show advertisements. These spaces which are usually located at the top or the side bar of the page, are called slots. Sponsored search deals with the way these advertisements are shown in the result page of an inquiry. Each time a keyword is searched by a user, the search engine requires to decide which advertisements to show among the different advertisers interested in that keyword. It also needs to determine the order and the location in which advertisements are shown. For this purpose, the search engine runs an online auction each time an inquiry takes place.

Each advertiser provides the search engine a bid, a list of relevant keywords and a maximum budget for a certain period of time. Using these data, the search engine conducts an online auction among advertisers interested in the searched keyword. It ranks the advertisers based on their bid and other parameters, and show them accordingly in the result page of the inquiry. Advertisers pay the search engines only when a user clicks on their advertisements. This payment scheme is called the "pay per click" auction.

The method used mainly in today's sponsored search auctions is the Generalized Second Price (GSP) auction. In this method, the search engine assigns each advertiser a weight based on her quality. The advertisers are scored by multiplying their weight and bid and are ranked based on their scores. The payment of each advertiser is the bid of the one who is right after her.

In GSP auctions, a Click Through Rate (CTR) is assigned to each advertiser which is the probability that a user clicks on a slot when it is occupied by the advertiser. Each advertiser has also a value which is defined as the expected profit for her if her advertisement is clicked on by a user. The utility of an advertiser is calculated as the subtraction of the value by the payment, multiplied by her CTR.

The assumption that an advertiser's CTR depends only on her slot and her quality is not realistic. Indeed, CTR also depends on the quality and position of the other advertisers who are displayed in the result page. Therefore, advertisers have externality effects on each other which influence their CTR and consequently affect their bidding behavior. Studying on these externality effects is the main focus of Chapter 3. In this chapter, first we review the studies on this topic and then we present the only experimental work on the externality effects by R. Gomes and N. Immorlica. They analyzed the clicking records associated to queries on some keywords in Microsoft's Live Search and showed the externality effects that advertisers have on each other. Later, we add a basic externality parameter to the basic model of sponsored search in particular and show its effects on the properties of the model.

Gathering search click data is extremely difficult due to the privacy issues of search engines. So in order to have a better understanding of the interaction among advertisers, we study the structure of the network of on-line advertisers in Chapter 4. At the beginning of this chapter, we discuss large-scale networks. One of the concepts that has been mostly studied in large-scale networks is the small-world property. A small-world network is a network which has a small average distance and high clustering coefficient. Of the three kinds of small-world networks, we focus on the ones that are scale-free. A scale-free network is a network whose degree distribution has power-law.

We present some examples of small-world scale-free networks such as the science collaboration network, the world wide web, the web of human sexual contacts and On-line Social Networks (OSNs). We discuss some of the earlier studies which show that they have a small-world scale-free structure. We also present Barabási-Albert model for scale-free networks which is a natural random process that makes a network with a power law degree distribution.

The main focus of this chapter is to analyze the structure of Google advertisement network which is a large-scale network. We chose Google because of its high revenue through online advertisement which is the number one among the different search engines. Google shows up to three advertisements on top of the result page and up to eight advertisements on the right side bar of the page. We show that this network is a small-world scale-free network. We model this network by a directed graph whose vertices are advertisers. On this graph an edge between two vertices is built if two advertisers have been shown on the same result page of an inquiry. If we number slots such that the three top slots are numbered top-down 1 to 3 and the eight right side slots are numbered 4 to 11, the direction of an edge in the graph is from the higher number vertex to the lower one.

The Google advertisement network has 81791 vertices and about 2 million edges. In this chapter, we describe how we collected the list of vertices and edges of this graph by using Python and C++ programming languages and the way we analyzed this graph using Pajek software (Pajek is a software for analyzing large-scale networks). Our results show that the Google advertisement network has the average distance of 2.88431 among reachable pairs, a high clustering coefficient of 0.5457577 and a power law degree distribution with degree exponent of close to 1.5. At the end of this chapter, we study the various cases of the Google advertisement network by applying some changes to the graph of the network.

1.1 Statement of Originality

The first part provides an alternative way to prove the Nikaidô-Sion version of the von Neumann minimax theorem based on elementary arguments. This approach, outlined by Ben-El-Mechaiekh and Dimand [10], could easily be adopted in a first course on Game Theory. We show in this thesis that the Klee-Berge intersection theorem, a key step in the elementary proof of the von Neumann theorem is equivalent to a "convex KKM Principle", thus extending to arbitrary topological vector spaces a result of Granas-Lassonde [23] shown to hold in super-reflexive Banach spaces.

The second part starts by extending a basic model for on-line auctions by adding a simple externality request. It shows that the extended model may not even have the basic properties of an ordinary auction. It also analyzes the structure of the underlying graph of Google advertisements network, and shows that it has a small average distance of 2.88431 and a high clustering coefficient of 0.5457577. It also shows that this network has Power-law degree distribution, and so it is a scale-free network.

Chapter 2

An elementary treatment of the Von Neumann minimax theorem and related results

2.1 Preliminaries

In this section we propose some preliminaries, definitions, examples and theorems needed for the sequel. A simple example of a two player zero-sum game with *pure* strategies is described by the pay-off matrix in Table 2.1.

	A	B	C
1	-3	-2	2
2	-1	0	4
3	-4	-3	1

Table 2.1: Pay-off matrix for player 1.

The payoff matrix for player 2 is the matrix above with the signs reversed. Here the strategy sets for both players are $X = \{1, 2, 3\}$ and $Y = \{A, B, C\}$, respectively. If player 1 chooses strategy 2 and player 2 chooses strategy C, when the payoff is allocated, player 1 gains 4 points and player 2 loses 4 points. If player 1 chooses strategy 1 or 3, she could gain at least -3 or -4 points respectively. While if she chooses strategy 2, she could gain at least -1 point. Therefore the secure strategy for player 1 is strategy 2. By choosing strategy 2 she can insure herself a minimal gain of -1 points, so $\alpha := \max_X \min_Y f(x, y) = -1$. Similarly, player 2 secures herself a gain of 1 point by choosing strategy A. Therefore $\beta := \min_Y \max_X f(x, y) = -1$. Thus, the value of this game is -1.

Another example of such games with *Mixed* strategies is given in Table 2.2. In this example the minimax choice for player 1 is strategy 2 since the worst possible result is then having to pay 1, while the simple minimax choice for player 2 is strategy B since the worst possible result is then no payment. Therefore this game does not have value with pure strategies.

	A	B	C
1	3	-2	2
2	-1	0	4
3	-4	-3	1

Table 2.2: Pay-off matrix for player 1.

In this example some choices are dominated by others and can be eliminated. For instance player 1 will not choose strategy 3 since either strategy 1 or 2 will produce a better result, no matter what player 2 chooses; player 2 will not choose strategy C since strategy A or B will produce a better result, no matter what player 1 chooses.

Player 1 can avoid having to make an expected payment of more than 1/3 by choosing strategy 1 with probability 1/6 and strategy 2 with probability 5/6, no matter what player 2 chooses. On the other side player 2 can ensure an expected gain of at least 1/3 by using a randomized strategy of choosing strategy A with probability 1/3 and strategy B with probability 2/3, no matter what player 1 chooses. Indeed by these mixed strategies this game has value of -1/3.

The starting point of our discussion is the simplest version of the theorem on the separation of convex sets in finite dimensions. We follow here the treatment by Magill and Quinzii [31]. We start with a preparatory lemma on the separation of a point and a convex set (which is in fact enough for our purposes!).

Lemma 2.1.1. ***(Separation of disjoint point and convex set)***

Let C be a nonempty closed convex subset of $\mathbb{R}^n$ and let $x \notin C, y$ denoting the projection of x onto C. The hyperplane H_u^y orthogonal to $u = x - y$ which passes through y strictly separates x and C, namely (see Figure 2.1),

$$u \cdot z \leq u \cdot y < u \cdot x, \ \forall z \in C.$$

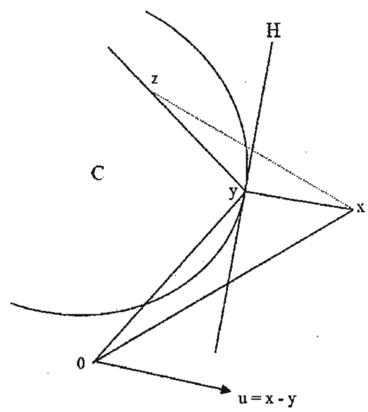

Figure 2.1: The hyperplane H orthogonal to u which passes through y, separates x from the convex set C.

Proof. Since C is closed and convex, the projection y of x onto C is unique (see e.g. [31]). Define, for a given $z \in C$, $\phi_z : [0,1] \to \mathbb{R}$ by:

$$\phi_z(t) = \|x - (tz + (1-t)y)\|^2,$$

The square of the distance between x and points on the line segment $[y, z] \subset C$. Az y is closest to x, ϕ_z attains its minimum on $[0,1]$ at $t = 0$ so that $\phi_z'(0) \geq 0, \forall z \in C$. Since

$$\phi_z = \|(x-y) + t(y-z)\|^2 = \|x-y\|^2 + 2t(x-y)\cdot(y-z) + t^2\|y-z\|^2,$$

we have:

$$\phi_z'(t) = 2t|y-z|^2 + 2(x-y)\cdot(y-z),$$

and $\phi_z'(0) = 2(x-y)\cdot(y-z)$. Thus $u\cdot(y-z) \geq 0$ for all $z \in C$, that is $u\cdot z \leq u\cdot y$. Since $u \cdot (x-y) = \|x-y\|^2$, $u \cdot (x-y) > 0$, that is $u \cdot y < u \cdot x$, and the proof is complete. □

This result extends to disjoint convex sets.

Theorem 2.1.1. ***(Separation of disjoint convex sets)***

(i) If K and C are non-empty convex subsets of $\mathbb{R}^n$ with $K \cap C = \emptyset$, then there exists $u \in \mathbb{R}^n, u \neq 0$, such that

$$\sup_{x \in C} u.x \leq \inf_{x' \in K} u.x' \tag{2.1}$$

(ii) If in addition K is compact and C is closed, then

$$\sup_{x \in C} u.x < \inf_{x' \in K} u.x'. \tag{2.2}$$

Proof. (i) Put $\tilde{C} = C - K = \{y \in \mathbb{R}^n \mid y = x - x', x \in C, x' \in K\}$. Since K and C are non-empty, $\tilde{C}$ is non-empty and convex as well. $0 \notin \tilde{C}$ as $K \cap C = \emptyset$. There are two cases to consider.

(a) $0 \notin \overline{\tilde{C}}$. By Lemma 2.1.1, there exists $u \in \mathbb{R}^n, u \neq 0$ such that $u.z < u.0 = 0, \forall z \in \overline{\tilde{C}}$. If we restrict ourself to the members of $\tilde{C}$ then, $u.x < u.x'$, $\forall x \in C$ and $\forall x' \in K$, which establishes (2.2).

(b) $0 \in \overline{\tilde{C}}$. Let $\{e^1, ..., e^m\}$ be a maximal collection of linearly independent vectors in $\tilde{C}$. Then for all $x \in \tilde{C}, \{x, e^1, ..., e^m\}$ are linearly dependent so that $x = \sum_{i=1}^m \lambda_i e^i$. We show that for all $\alpha > 0, -\alpha \sum_{i=1}^m e^i \notin \overline{\tilde{C}}$.

Suppose there exists $\alpha > 0$ such that $-\alpha \sum_{i=1}^m e^i \in \overline{\tilde{C}}$. Then there exists a sequence $\{x^\nu = \sum_{i=1}^m \lambda_i^\nu e^i\} \in \tilde{C}$ such that $x^\nu \to -\alpha \sum_{i=1}^m e^i$. since $\{e^1, ..., e^m\}$ are linearly independent, $\lambda_i^\nu \to -\alpha, i = 1, ..., m$, so that for sufficiently large $\nu, \lambda_i^\nu < 0, i = 1, ..., m$. Then we have $\frac{1}{1-\sum_{i=1}^m \lambda_i^\nu} \in [0,1]$, and by the convexity of $\tilde{C}$:

$$\left(\frac{1}{1-\sum_{i=1}^m \lambda_i^\nu}\right) x^\nu + \left(1 - \frac{1}{1-\sum_{i=1}^m \lambda_i^\nu}\right) e^i \in \tilde{C},$$

which is:
$$0 = \left(\frac{1}{1-\sum_{i=1}^m \lambda_i^\nu}\right)x^\nu - \sum_{i=1}^m \left(\frac{\lambda_i^\nu}{1-\sum_{i=1}^m \lambda_i^\nu}\right)e^i \in \tilde{C},$$
contradicting $0 \notin \tilde{C}$. Thus for all $k \in \mathbb{N}, -\frac{1}{k}\sum_{i=1}^m e^i \notin \overline{\tilde{C}}$. By Lemma 2.1.1, there exists $u_k \in \mathbb{R}^n$, which can be chosen with $\|u_k\| = 1$, such that:
$$u_k.z < u_k.\left(-\frac{1}{k}\sum_{i=1}^m e^i\right), \forall z \in \overline{\tilde{C}}.$$

Choosing a subsequence such that $u_k \to u$ and restricting the inequality to elements of $\tilde{C}$ gives:
$$u.z \leq 0, \ \forall z \in \tilde{C} \Leftrightarrow u.x \leq u.x', \ \forall x \in C, \ \forall x' \in K,$$
which proves (2.1).

(ii) If in addition K is compact and C is closed, then $\tilde{C}$ is closed (the sum of a closed set and a compact set is closed). Thus $0 \notin \tilde{C} = \overline{\tilde{C}}$, and the first case of the proof of (i) also holds. □

We draw the reader's attention that part (ii) is what is needed in the proof of *Implication 1* in Figure 1.1. The proof of (ii) is an immediate consequence of Lemma 2.1.1.

Throughout this chapter, we work in the framework of general real topological vector spaces.

Definition 2.1.1. *[38] Given a vector space L over $\mathbb{R}$ and a topology τ on $\mathbb{R}$ the pair (L, τ) is called a* Topological Vector Space over $\mathbb{R}$ *if these two conditions are satisfied:*
$$(x, y) \to x + y \ \textit{ is continuous on } \ \mathbb{R} \times \mathbb{R} \ \textit{ into } \ \mathbb{R}.$$
$$(\lambda, x) \to \lambda x \ \textit{ is continuous on } \ \mathbb{R} \times \mathbb{R} \ \textit{ into } \ \mathbb{R}.$$

Throughout this chapter, topological vector spaces are supposed to be Hausdorff (i.e. distinct points can be separated by disjoint open neighborhoods). We shall also consider numerical functions that may be short of being fully continuous or convex. Indeed, while von Neumann's original formulation of the minimax theorem considered linear forms, he became quickly aware that the convexity of sub-level sets of the objective functions was the key geometric consideration for the existence of a saddle point. He thus alluded to the concept of quasi-convexity as described below.

Definition 2.1.2. *A real function $f : X \to \mathbb{R}$ defined on a subset X of a topological vector space is:*
(i) quasiconvex *if $\forall \lambda \in \mathbb{R}$, the level set $\{x \in X : f(x) < \lambda\}$ is a convex subset of X.*
(ii) upper semicontinuous (u.s.c.) *if $\forall \lambda \in \mathbb{R}$, the level set $\{x \in X : f(x) < \lambda\}$ is an open subset of X.*

A function f is *quasiconcave* if $-f$ is quasiconvex; it is *lower semicontinuous (l.s.c.)* if $-f$ is u.s.c. Note that f is quasiconvex on X if and only if:

$$f(\mu x_1 + (1-\mu)x_2) \leq max\{f(x_1), f(x_2)\} \text{ for all } x_1, x_2 \in X \text{ and all } \mu \in [0,1]$$

Convex functions are clearly quasiconvex (see Crouzeix [27] for a comparative study of quasiconvexity).

Example 2.1.1. *The function $f : \mathbb{R}^+ \to \mathbb{R}$ with $f(x) = \ln x$ or the function shown in Figure 2.2 are quasiconvex functions which are not convex.*

Example 2.1.2. *Let f be defined on $X = [0,1] \in \mathbb{R}$ by:*

$$f(x) = \begin{cases} 1, & \textit{for } 0 \leq x \leq 1/3 \\ 0, & \textit{for } 1/3 < x \leq 1 \end{cases}$$

Then f is upper semicontinuous on X and is clearly not continuous.

Remark 2.1.1. *(i) The extreme value theorem does not require full continuity of the objective function. Indeed, lower semicontinuous functions on compact*

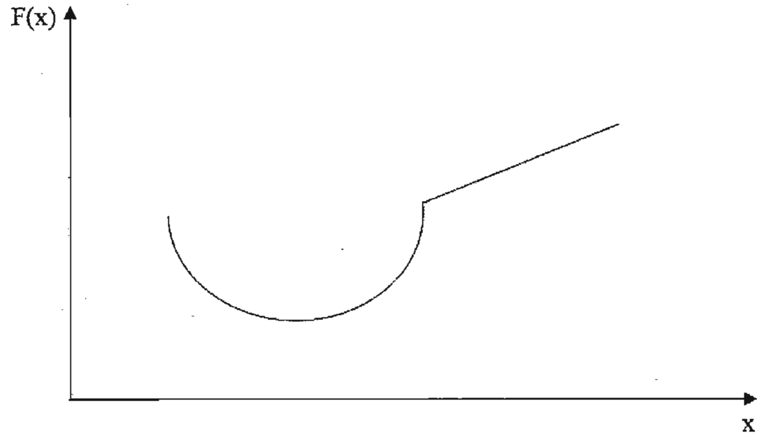

Figure 2.2: Example of quasiconvex function.

domains achieve their minimum, namely: If $f(x)$ is a l.s.c. function defined on a compact set X, then it achieves its minimum:

$$\exists \bar{x} \in X, \qquad f(\bar{x}) = \min_{x \in X} f(x)$$

(ii) The upper envelope of l.s.c. function is also l.s.c., more precisely: if $\{f_i\}_{i \in I}$ is a family of l.s.c. functions, the $\sup_{i \in I} f_i$ is also an l.s.c. function.

Remark 2.1.2. *Given two sets X and Y, a set-valued map is a transformation F that assigns to each element $x \in X$, a subset $F(x)$ of Y.*

In 1929 Knaster-Kuratowski-Mazurkiewicz using Sperner's lemma as a tool, established the following result on the non-empty intersection of a family of closed subsets of a Euclidean space [29]:

Theorem 2.1.2. ***(KKM lemma)***
Suppose that a simplex S^m is covered by the closed sets C_i for $i \in I = \{1, ..., m\}$ and that for all $I_k \subset I$ the face of S that is spanned by e_i, for $i \in I_k$, is covered by C_i. Then all the C_i have a common intersection point.

In 1978 Dugundji-Granas [16] introduced the following class of set-valued maps:

Definition 2.1.3. *Given an arbitrary subset X of a real vector space L, a set-valued map $\Gamma : X \to 2^E$ is said to be a* KKM map *if for every finite subset $A : \{x_1, ..., x_n\} \subseteq X$:*

$$Conv(A) \subset \bigcup_{i=1}^{n} \Gamma(x_i).$$

Ky Fan in 1961 [17] extended the KKM lemma to topological vector spaces of any dimension. Ky Fan's extension, with this terminology, asserts that, given a KKM closed valued map Γ defined on a subset X of a topological vector space, the family $\Gamma(x) : x \in X$ has the finite intersection property. An additional suitable compactness condition, implies the non-emptiness of the intersection for the entire family (see Theorem 2.2.6). This principle is the basic ingredient in the proofs of "intersection" theorems and related fixed point theorems of topological nature (including the famous Brouwer fixed point theorem). There have subsequently been numerous extensions of the KKM principle with the aims of replacing linear convexity by topological substitutes (see e.g., [24]) or by abstract convex structures.

The KKM principle has a "dual" formulation as a fixed point theorem for a set-valued map. Indeed, as we will see, a KKM map gives rise to a so-called Ky Fan mapping (see e.g. [9] and references there). Those maps are extended to:

Definition 2.1.4. *[9] A set-valued map $A : X \to 2^Y$ is a Φ-map (written $A \in \Phi(X, Y)$) if and only if:*

(i) $A^{-1}(y)$ is convex in X for all $y \in Y$;

(ii) A has a set-valued selection with open values and nonempty fibers, i.e., $A(x) \supseteq \widetilde{A}(x)$ open in Y for all $x \in X$, and $\widetilde{A}^{-1}(y) \neq \emptyset$ for all $y \in Y$.

A map $B : X \to 2^Y$ is a Φ*-map if and only if its inverse $B^{-1} : Y \to 2^X$ is a Φ-map.

Definition 2.1.5. *A family $\{\Gamma(x)\}_{x \in X}$ of sets defined on an arbitrary set X is said to have the* finite intersection property *if and only if for every finite subset $\{x_1, ..., x_n\} \subset X$, $\bigcap_{i=1}^{n} \Gamma(x_i) \neq \emptyset$.*

It is well-known that a topological vector space Y is compact if and only if for every collection $\{\Gamma(x)\}_{x\in X}$ of closed sets having the finite intersection property, $\bigcap_{x\in X} \Gamma(x) \neq \emptyset$.

We end this section by recalling the basic concept of a partition of unity subordinated to a cover.

Definition 2.1.6. *[4] A* partition of unity *on a set X is a family $\{f_i\}_{i\in I}$ of functions from X into [0,1] such that at each $x \in X$, only finitely many functions in the family are nonzero and*

$$\sum_{i\in I} f_i(x) = 1,$$

where, by convention, the sum of an arbitrary collection of zero is zero.

A partition of unity is *subordinated* to a cover U of X if each function vanishes outside some member of U. For a topological space, a partition of unity is called *continuous* if each function is continuous. Every open cover of a paracompact (respectively, compact) topological space has a locally finite (finite, respectively) subordinated partition of unity.

2.2 The Von Neumann's Minimax Principles and Related Results from Convex Analysis

We shall proceed in two steps to show the solid line implications in Figure 1.1. We start first with the implications 1 to 4 and will follow with the remaining implications.

2.2.1 The Elementary Proof

We are now ready to detail the passage from the separation theorem to the Nikaidô-Sion version of the minimax theorem. The first step is the intersection theorem of Berge-Klee.

Theorem 2.2.1. *[13, 28] **(Berge-Klee Theorem)***

Let C and $C_1, ..., C_n$ be closed convex sets in a Euclidean space satisfying:

(i) $C \cap \bigcap_{i=1, i\neq j}^{n} C_i \neq \emptyset$ for $j = 1, 2, ..., n$;

(ii) $C \cap \bigcap_{i=1}^{n} C_i = \emptyset$.

Then $C \not\subseteq \bigcup_{i=1}^{n} C_i$.

Proof. We follow Victor Klee's proof. We may assume that C and C_i's are compact (otherwise, we replace C by the convex finite polytope $C' := Conv\{y_j : j = 1, .., n\}$, where $y_j \in C \cap \bigcap_{i=1, i\neq j}^{n} C_i$, and C_i by $C_i' := C_i \cap C'$). To prove this theorem we use induction on n.

Step 1: If $n = 1$, by (i) C is nonempty, and by (ii), C and C_1 are disjoint. Thus clearly $C \not\subseteq C_1$.

Step 2: We suppose that the theorem holds for $n = k - 1$.

Step 3: Assume $n = k$ i.e. $\{C, C_1, ..., C_k\}$ is a collection of compact convex sets, with $C \cap \bigcap_{i=1, i\neq j}^{k} C_i \neq \emptyset$ for any $j = 1, 2, ..., k$, and $C \cap \bigcap_{i=1}^{k} C_i = \emptyset$. We want to show that $C \not\subseteq \bigcup_{i=1}^{k} C_i$.

We have $(C \cap C_k) \cap \bigcap_{i=1}^{k-1} C_i = C \cap \bigcap_{i=1}^{k} C_i = \emptyset$. As $C \cap C_k$ and $\bigcap_{i=1}^{k-1} C_i$ are nonempty convex subset of a Euclidean space, by Theorem 2.1.1 they can be strictly separated by a hyperplane H. Thus $H \cap \bigcap_{i=1}^{k-1} C_i = \emptyset$, and hence $(H \cap C) \cap (\bigcap_{i=1}^{k-1} C_i \cap H) = \emptyset$. Putting $C' := H \cap C$ and $C_i' := H \cap C_i$, we have $C' \cap (\bigcap_{i=1}^{k-1} C_i') = \emptyset$.

From $C \cap \bigcap_{i=1, i\neq j}^{k} C_i \neq \emptyset$ for $j = 1, 2, ..., k$, we have $\exists y_0 \in C \cap \bigcap_{i=1, i\neq j}^{k} C_i \neq \emptyset$ for $j = 1, 2, ..., k-1$ which means $y_0 \in C \cap C_k$ and $\exists y_k \in C \cap \bigcap_{i=1}^{k-1} C_i \neq \emptyset$, that is $y_k \in \bigcap_{i=1}^{k-1} C_i$. Therefore y_0 and y_k are separated by H. Let $\overline{z} = [y_0, y_k] \cap H$. Then $\overline{z} \in C \cap H = c'$ and $\overline{z} \in (\bigcap_{i=1, i\neq j}^{k-1} C_i) \cap H = \bigcap_{i=1, i\neq j}^{k-1} C_i'$ for any $j = 1, 2, ..., k - 1$. Therefore $C' \cap \bigcap_{i=1, i\neq j}^{k-1} C_i' \neq \emptyset$ for any $j = 1, 2, ..., k - 1$.

Using step 2 of the induction, we have $C' \not\subseteq \bigcup_{i=1}^{k-1} C'_i$ which means $C \cap H \not\subseteq \bigcup_{i=1}^{k-1} (C_i \cap H)$. As $(H \cap C) \cap C_k = \emptyset$, we have $C \cap H \not\subseteq \bigcup_{i=1}^{k} (C_i \cap H)$, which gives $C \not\subseteq \bigcup_{i=1}^{k} C_i$. □

We now show that the Berge-Klee theorem yields a nonlinear alternative for systems of inequalities, the key ingredient in the proof of the minimax theorem.

Theorem 2.2.2. *[10] **(A First Alternative for Systems of Nonlinear Inequalities)***

Let X and Y be two convex subsets of topological vector spaces, with Y compact,and let $f_1, f_2, f_3, f_4 : X \times Y \to \mathbb{R}$ be four functions satisfying:

(i) $f_1(x,y) \leq f_2(x,y) \leq f_3(x,y) \leq f_4(x,y)$ for all $(x,y) \in X \times Y$;

(ii) $y \mapsto f_1(x,y)$ is lower semicontinuous and quasiconvex on Y for each fixed $x \in X$;

(iii) $x \mapsto f_2(x,y)$ is quasiconcave on X for each fixed $y \in Y$;

(iv) $y \mapsto f_3(x,y)$ is quasiconvex on Y for each fixed $x \in X$;

(v) $x \mapsto f_4(x,y)$ is upper semicontinuous and quasiconcave on X for each fixed $y \in Y$;

Then for any $\lambda \in \mathbb{R}$, the following alternative holds:

(A) there exists $\overline{x} \in X$ such that $f_4(\overline{x}, y) \geq \lambda$, for all $y \in Y$; or

(B) there exists $\overline{y} \in Y$ such that $f_1(x, \overline{y}) \leq \lambda$, for all $x \in X$

Proof. Suppose that the nonlinear alternative does not hold, i.e. both (A) and (B) fail. If (A) fails then $\forall x \; \exists \overline{y} \; s.t \; f(x, \overline{y}) < \lambda$ which implies that the collection of open level sets $\{U_y := \{x \in X : f_4(x,y) < \lambda\} : y \in Y\}$ is a cover of X (each U_y is open because $f_4(x,y)$ is u.s.c). Similarly, if (B) fails, the collection $\{V_x := \{y \in Y : f_1(x,y) > \lambda\} : x \in X\}$ is an open cover of Y.

Since Y is compact, $\{V_x : x \in X\}$ admits a finite subcover $\{V_{x_k} : k = 1, ..., m\}$. Let $C := Conv\{x_k : k = 1, ...m\}$. As C is compact and in X, it can be covered by a finite subcollection $\{U_{y_i} : i = 1, ..., n\}$. We consider a minimal cover $\{U_{y_i}\}$, in the sense that $C \subseteq \bigcup_{i=1}^{n} U_{y_i}$ but $C \nsubseteq \bigcup_{i=1, i\neq j}^{n} U_{y_i}$ for $j = 1, ..., n$.

For $i = 1, ..., n$, let $C_i := \{x \in X : f_4(x, y_i) \geq \lambda\}$. Since f_4 is u.s.c and quasiconcave, C_i's are closed convex subsets of X. The fact that C is covered by $\{U_{y_i}\}$ implies the emptiness of the intersection $C \cap \bigcap_{i=1}^{n} C_i$. The minimality of $\{U_{y_i}\}$ causes $C \cap \bigcap_{i=1, i\neq j}^{n} C_i \neq \emptyset$ for $j = 1, ..., n$. By Theorem 2.2.1, $C \nsubseteq \bigcup_{i=1}^{n} C_i$, which implies the existence of $x_0 \in C$ with $x_0 \notin \bigcup_{i=1}^{n} C_i$, that is $x_0 \notin C_i$ thus $f_4(x_0, y_i) < \lambda$, hence $f_3(x_0, y_i) < \lambda$ for $i = 1, ..., n$. The quasiconvexity of $f_3(x_0, .)$ implies that

$$\exists x_0 \in C \text{ such that } f_3(x_0, y) < \lambda, \forall y \in D := Conv\{y_i : i = 1, ..., n\},$$

similarly

$$\exists y_0 \in D \text{ such that } f_2(x, y_0) > \lambda, \forall x \in C,$$

this yields

$$\lambda < f_2(x_0, y_0) \leq f_3(x_0, y_0) < \lambda,$$

a contradiction. □

This nonlinear alternative yields:

Theorem 2.2.3. ***(InfSup Theorem)***

Let X and Y be two convex subsets of topological vector spaces, with Y compact, and let $f_1, f_2, f_3, f_4 : X \times Y \to \mathbb{R}$ be four functions satisfying:

(i) $f_1(x, y) \leq f_2(x, y) \leq f_3(x, y) \leq f_4(x, y)$ for all $(x, y) \in X \times Y$;

(ii) $y \mapsto f_1(x, y)$ is lower semicontinuous and quasiconvex on Y for each fixed $x \in X$;

(iii) $x \mapsto f_2(x, y)$ is quasiconcave on X for each fixed $y \in Y$;

(iv) $y \mapsto f_3(x, y)$ is quasiconvex on Y for each fixed $x \in X$;

(v) $x \mapsto f_4(x, y)$ is upper semicontinuous and quasiconcave on X for each fixed

$y \in Y$.

Then:

$$\alpha = \sup_X \inf_Y f_4(x,y) \geq \min_Y \sup_X f_1(x,y) = \beta.$$

Proof. Assume $\alpha < \beta$, then $\exists \lambda \in \mathbb{R}, \alpha < \lambda < \beta$. By Theorem 2.2.2 (A) or (B) should hold.

If (A) holds, there exists $\overline{x} \in X$ such that $f_4(\overline{x}, y) \geq \lambda$, for all $y \in Y$, hence $\inf_Y f_4(\overline{x}, y) \geq \lambda$. Then $\alpha = \sup_X \inf_Y f_4(x,y) \geq \lambda$ which is a contradiction.

If (B) holds, there exists $\overline{y} \in Y$ such that $f_1(x, \overline{y}) \leq \lambda$, for all $x \in X$, hence $\sup_X f_1(x, \overline{y}) \leq \lambda$. Then (by Remark 2.1.1 the minimum exists) $\beta = \min_Y \sup_X f_1(x,y) \leq \lambda$ which is a contradiction.

Hence neither (A) nor (B) of Theorem 2.2.2 can hold. This is absurd, so $\alpha \geq \beta$. □

Putting $f_1 = f_2 = f_3 = f_4 = f$ in Theorem 2.2.3, we obtain the Nikaidô-Sion (see [35, 39]) version of the minimax theorem of von Neumann.

Theorem 2.2.4. ***(Nikaidô-Sion formulation of the minimax theorem of Von Neumann)***

Let X and Y be convex subsets of topological vector spaces, with Y compact, and let f be a real function on $X \times Y$ such that:

(i) $y \mapsto f(x,y)$ is lower semicontinuous and quasiconvex on Y for each fixed $x \in X$;

(ii) $x \mapsto f(x,y)$ is upper semicontinuous and quasiconcave on X for each fixed $y \in Y$;

Then

$$\sup_X \min_Y f(x,y) = \min_Y \sup_X f(x,y).$$

Proof. Since $\sup_X \min_Y f(x,y) \leq \min_Y \sup_X f(x,y)$ is always true and Theorem 2.2.3 implies that $\sup_X \min_Y f(x,y) \geq \min_Y \sup_X f(x,y)$, we have $\sup_X \min_Y f(x,y) = \min_Y \sup_X f(x,y)$. □

It is crucial to point out, here, that this first alternative for systems of nonlinear inequalities is less general than the one presented first in [9] and used below. But it is more than adequate for deriving the minimax theorem. The central difference between the two alternatives resides in the additional convexity assumptions in hypotheses (ii) and (v). These very quasiconvexity/concavity assumptions on f_1 and f_4 are what give the first alternative its "elementary" character. We do not see, at this moment, if it is possible to derive the second alternative from the first. Doing this, as the remainder will show, would amount to providing the elementary proof of the KKM Principle, and hence of the Brouwer fixed point theorem.

We now continue with the treatment of the lower loop and the lateral arrows in Figure 1.1.

2.2.2 Intersection Theorems, Fixed Point, and Coincidence

The starting point is the following version of the KKM principle due to Ky Fan:

Theorem 2.2.5. *[17] (**KKM Principle of Ky Fan**)*
Let $\Gamma : X \to 2^Y$ be a set-valued map and X and Y be subsets of real topological vector space E, with Γ a KKM map and $\Gamma(x)$ closed in Y for all $x \in X$. Then $\{\Gamma(x)\}_{x\in X}$ has the finite intersection property.
Assume that any one of the following four conditions holds:
(a) Y compact;
(b) each $\Gamma(x)$ compact;
(c) there exists $x_0 \in X$ with $\Gamma(x_0)$ compact;
(d) there exists $x_1, ..., x_n$ in X with $\bigcap_{i=1}^{n} \Gamma(x_i)$ compact.
Then $\bigcap_{x\in X} \Gamma(x) \neq \emptyset$.

Remark 2.2.1. *Note that the compactness conditions above are increasing in generality, i.e. (a) $\Rightarrow$ (b) $\Rightarrow$ (c) $\Rightarrow$ (d). Ky Fan considered a yet more general compactness condition in [19] which is sufficient to reduce the problem to finite dimension, namely: there exists X_0 in X such that $X_0 \subset C$ which C is a convex* *compact set in* X *and* $\bigcap_{x\in X_0} \Gamma(x) = K$ *with* K *compact in* Y. *A dual formulation of this condition was used in [9] in the context of fixed points for set-valued maps to relax the compactness of the domains by a control outside of a compact subset.*

The KKM Principle of Ky Fan can be generalized to involve two mappings:

Theorem 2.2.6. *[16] **(Dugundji-Granas KKM Theorem)***
Let $\Gamma, \tilde{\Gamma} : X \to 2^E$ *be two set-valued maps defined on an arbitrary subset* X *of a real topological vector space* E*, satisfying:*
(i) Γ *is a selection of* $\tilde{\Gamma}$ *(i.e,* $\Gamma(x) \subseteq \tilde{\Gamma}(x)$ *for all* $x \in X$;
(ii) Γ *is a KKM-map;*
(iii) all values of $\tilde{\Gamma}$ *are closed;*
(iv) $\bigcap_{x\in X} \tilde{\Gamma}(x) \neq \emptyset \Rightarrow \bigcap_{x\in X} \Gamma(x) \neq \emptyset$.
Then $\{\Gamma(x)\}_{x\in X}$ *has the finite intersection property.*
If any one of following four conditions holds:
(a) Y *compact;*
(b) each $\tilde{\Gamma}(x)$ *compact;*
(c) there exists $x_0 \in X$ *with* $\tilde{\Gamma}(x_0)$ *compact;*
(d) there exists $x_1, ..., x_n$ *in* X *with* $\bigcap_{i=1}^{n} \Gamma(x_i)$ *compact.*
Then $\bigcap_{x\in X} \Gamma(x) \neq \emptyset$.

This version of the KKM Principle can be phrased in analytical terms as another alternative for systems of nonlinear inequalities (see e.g., [9]).

Theorem 2.2.7. ***(A Second Alternative for Systems of Nonlinear Inequalities)***
Let X *and* Y *be two convex subsets of topological vector spaces, with* Y *compact,and let* $f_1, f_2, f_3, f_4 : X \times Y \to \mathbb{R}$ *be four functions satisfying:*
(i) $f_1(x,y) \leq f_2(x,y) \leq f_3(x,y) \leq f_4(x,y)$ *for all* $(x,y) \in X \times Y$;
(ii) $y \mapsto f_1(x,y)$ *is lower semicontinuous on* Y *for each fixed* $x \in X$;
(iii) $x \mapsto f_2(x,y)$ *is quasiconcave on* X *for each fixed* $y \in Y$;

(iv) $y \mapsto f_3(x,y)$ *is quasiconvex on* Y *for each fixed* $x \in X$;

(v) $x \mapsto f_4(x,y)$ *is upper semicontinuous on* X *for each fixed* $y \in Y$;

Then for any $\lambda \in \mathbb{R}$, *the following alternative holds:*

(A) there exists $\overline{x} \in X$ *such that* $f_4(\overline{x}, y) \geq \lambda$, *for all* $y \in Y$; *or*

(B) there exists $\overline{y} \in Y$ *such that* $f_1(x, \overline{y}) \leq \lambda$, *for all* $x \in X$

Proof. First step. we show that if this theorem holds with X,Y both convex compact, then it holds with X,Y both convex but only Y compact. Assume that alternative (B) fails, then $\forall y \in Y, \exists x \in X$ with $f_1(x,y) > \lambda$. So $\{\{y \in Y : f_1(x,y) > \lambda\}\}_{x \in X}$ is an open cover of Y. Since Y is compact, it has a finite subcover $\{y \in Y : f_1(x_i, y) > \lambda\}_{i=1}^n$. Put $C = Conv\{x_1, ..., x_n\} \subset X$. Since both C and Y are convex compact, we will have at least one of the following: (a) there exists $\overline{x} \in C$ such that $f_4(\overline{x}, y) \geq \lambda$, for all $y \in Y$, or (b) there exists $\overline{y} \in Y$ such that $f_1(x, \overline{y}) \leq \lambda$, for all $x \in C$.

We show that (b) fails: Assume (b) holds, then $f_1(x, \overline{y}) \leq \lambda$ for all $i = 1, ...n$, which contradict with the fact that $Y = \bigcup_{i=1}^{n} \{y \in Y : f_1(x_i, y) > \lambda\}$. Therefore (a) holds, hence alternative (A) holds.

We showed that if alternative (B) fails then (A) holds. Similarly we have the converse.

Second step. We prove this theorem by assuming both X and Y are convex compact. Define set-valued maps $F, \tilde{F} : X \to 2^Y$ with:

$$F(x) = \{y \in Y : f_2(x,y) > \lambda\},$$

and

$$\tilde{F}(x) = \{y \in Y : f_1(x,y) > \lambda\}.$$

Also define set-valued maps $G, \tilde{G} : Y \to 2^X$ with:

$$G(y) = \{x \in X : f_3(x,y) < \lambda\},$$

and

$$\tilde{G}(y) = \{x \in X : f_4(x,y) < \lambda\}.$$

Clearly $\tilde{F}(x) \subseteq F(x)$ for all $x \in X$, and $\tilde{G}(y) \subseteq G(y)$ for all $y \in Y$. If $\tilde{F}$ is not surjective then (B) holds, similarly if $\tilde{G}$ is not surjective then (A) holds.

We complete the proof by showing that $\tilde{F}$ and $\tilde{G}$ can not be both surjective: Assume $\tilde{F}$ and $\tilde{G}$ are both surjective, and define $\Gamma, \tilde{\Gamma} : X \times Y \to 2^{X \times Y}$ by

$$\Gamma(x,y) = X \times Y \setminus (G(y) \times F(x)),$$

and

$$\tilde{\Gamma}(x,y) = X \times Y \setminus (\tilde{G}(y) \times \tilde{F}(x)).$$

Bing $\tilde{F}$ and $\tilde{G}$ both surjective implies that F and G are surjective. Observe that $\bigcap\limits_{(x,y)\in X\times Y} \Gamma(x,y) = \emptyset$. Therefore the conclusion of the Theorem 2.2.6 does not hold for Γ and $\tilde{\Gamma}$. Clearly $\Gamma(x,y) \subseteq \tilde{\Gamma}(x,y)$ for all $(x,y) \in X \times Y$. Also $\tilde{\Gamma}(x,y)$ is closed due to the lower semicontinuity of $\tilde{f}$ and upper semicontinuity of $\tilde{g}$. As $\tilde{F}$ and $\tilde{G}$ are surjective $\bigcap\limits_{(x,y)\in X\times Y} \tilde{\Gamma}(x,y) = \emptyset$.

Hypothesis (i),(iii) and (iv) of Theorem 2.2.6 holds, but not the conclusion. This implies that Γ is not a KKM map. It means there exists a convex combination $(x_0, y_0) = \sum\limits_{i=1}^{n} \lambda_i (x_i, y_i)$ with
$(x_0, y_0) \notin \bigcup\limits_{i=1}^{n} \Gamma(x_i, y_i) \Leftrightarrow (x_0, y_0) \in \bigcap\limits_{i=1}^{n} (G(y_i) \times F(x_i)) \Leftrightarrow (y_i, x_i) \in G^{-1}(x_0) \times F^{-1}(y_0)$ for all $i \in \{1, ..., n\}$.
Since $G^{-1}(x_0) \times F^{-1}(y_0)$ is a convex subset of $Y \times X$, it follows that $(y_0, x_0) \in G^{-1}(x_0) \times F^{-1}(y_0)$ that is $f_3(x_0, y_0) < \lambda < f_2(x_0, y_0)$. This contradicts $f_3(x_0, y_0) \geq f_2(x_0, y_0)$. □

Note that this second alternative is more general than the first one (Theorem 2.2.2) as the convexity assumptions on the first and last functions are removed. Such alternatives can be expressed in geometric terms as coincidence theorems for set-valued maps.

Theorem 2.2.8. *[9] (A Coincidence Principle)*
Let $A \in \Phi(X, Y)$ and $B \in \Phi^*(X, Y)$ where X and Y are nonempty convex subsets

of topological vector spaces. If Y is compact, then there exists $(x_0, y_0) \in X \times Y$ such that

$$y_0 \in A(x_0) \cap B(x_0).$$

Proof. Assume $A \in \Phi(X,Y)$, $B \in \Phi^*(X,Y)$, and X, Y are nonempty convex subsets of topological vector spaces. We define $\tilde{f}, f, g, \tilde{g} : X \times Y \to \mathbb{R}$ as:

$$\tilde{f}(x,y) = \begin{cases} 0, & \text{if } y \notin \tilde{A}(x) \\ 1, & \text{if } y \in \tilde{A}(x) \end{cases} \qquad f(x,y) = \begin{cases} 0, & \text{if } y \notin A(x) \\ 1, & \text{if } y \in A(x) \end{cases}$$

$$g(x,y) = \begin{cases} 0, & \text{if } y \in B(x) \\ 1, & \text{if } y \notin B(x) \end{cases} \qquad \tilde{g}(x,y) = \begin{cases} 0, & \text{if } y \in \tilde{B}(x) \\ 1, & \text{if } y \notin \tilde{B}(x) \end{cases}$$

For each fixed $x \in X$ and each $\lambda \in \mathbb{R}$, $S = \{y | \tilde{f}(x,y) > \lambda\}$ is an open set in Y. If $\lambda \geq 1$ then $S = \emptyset$ which is open, if $0 \leq \lambda < 1$ then $S = \tilde{A}(x)$ which is open in Y, and finally if $\lambda < 0$, $S = Y$ which is open. Therefore $\tilde{f}$ is l.s.c. on Y for each fixed $x \in X$. With the same argument f is quasiconcave on X for each fixed $y \in Y$, g is quasiconvex on Y for each fixed $x \in X$, and $\tilde{g}$ is u.s.c. on X for each fixed $y \in Y$. Hence, the conditions (ii) to (v) of the second alternative theorem are satisfied. Since $\tilde{B}(x) \neq \emptyset$ for all $x \in X$, and $\tilde{A}^{-1}(y) \neq \emptyset$ for all $y \in Y$, both (A) and (B) of Theorem 2.2.7 fail. All the conditions except the first one of the second alternative theorem hold, but its conclusion does not hold. Therefore the first condition is not satisfied.

Then, since clearly $\tilde{f}(x,y) \leq f(x,y)$ and $g(x,y) \leq \tilde{g}(x,y)$, it follows that $f(x,y) \leq g(x,y)$ fails, That is:

$$0 = g(x_0, y_0) < f(x_0, y_0) = 1 \text{ for some } (x_0, y_0) \in X \times Y,$$

which is equivalent to $y_0 \in A(x_0) \cap B(x_0)$. □

The coincidence theorem has many applications in the theory of games and other areas such as minimax inequality, fixed point theory, existence of minimizable quasiconvex functions, etc ... (see e.g., [15, 16, 9] for related results). It is known that it implies the existence of Nash equilibria for generalized games with n-players involving objective functions verifying adequate quasiconvexity/concavity and continuity assumptions (see e.g. [16]). This coincidence theorem is in fact a by-product of the Browder-Fan fixed point theorem. We chose here to provide the proof of the converse, namely implication (7) in Figure 1.1. We require a preliminary result saying that every continuous function f on a compact set admits a $\mathbf{\Phi}^*-$majorant that is within an $\epsilon-$tubular neighborhood of the graph of f. More precisely,

Lemma 2.2.1. *[12] Given a continuous function $f : C \longrightarrow Y$ from a compact metric space C into a convex subset Y of a normed space, for any $\epsilon > 0$, there exists a $\mathbf{\Phi}^*-$map $\Phi_\epsilon : C \to 2^Y$ such that:*

$$f(x) \in \Phi_\epsilon(x) \subset B_\epsilon(f(B_\epsilon(x))), \forall x \in C,$$

where, for a given set Z, $B_\epsilon(Z)$ denotes the $\epsilon-$open ball around Z.

Proof. By continuity, $\forall x \in C, \exists 0 < \delta_x < \epsilon$ with $f(B_{\delta_x}(x) \cap C) \subset B_\epsilon(f(x)) \cap Y$. Let $B := \{B_{\delta_{x_i}}(x_i) \cap C : i = 1, ..., p\}$ be a finite cover of the compact set C, and for each $x \in C$, let $I(x) := \{i \in \{1, ..., p\} : x \in B_{\delta_{x_i}}(x_i)\}$ be the set of essential indices of x w.r.t. the open cover B. Define the map Φ_ϵ by setting:

$$\Phi_\epsilon(x) := \bigcap_{i \in I(x)} (B_\epsilon(f(x_i)) \cap Y), \forall x \in C.$$

Clearly, the values of Φ_ϵ are convex non-empty and its graph is open (indeed, for each $x \in C$, the open set $\bigcap_{i \in I(x)}(B_{\delta_{x_i}}(x_i) \cap C) \times \Phi_\epsilon(x)$ is a neighborhood of $\{x\} \times \Phi_\epsilon(x)$), thus Φ_ϵ has open fibers, thus $\Phi_\epsilon \in \mathbf{\Phi}^*(C, Y)$. By definition, $f(x) \in \Phi_\epsilon(x) \subset B_\epsilon(f(B_\epsilon(x))), \forall x \in C$. □

Now we are ready to prove a generalization of The Browder-Ky Fan Fixed Point Theorem due to Ben-El-Mechaiekh et al. [9].

Theorem 2.2.9. ***(The Browder-Ky Fan Fixed Point Theorem)***
If $A \in \mathbf{\Phi}^*(X)$ (equivalently $A \in \mathbf{\Phi}(X)$) where X is compact convex, in a topological vector space, then A has a fixed point $x_0 \in A(x_0)$.

Proof. First, we reduce the Browder-Ky Fan fixed point theorem to finite dimensions by using a standard selection argument (see [9]).

Indeed, given a map $A \in \mathbf{\Phi}^*(X)$, where X is compact convex, in a topological vector space, the fact that $A(x) \neq \emptyset$ for all $x \in X$, implies that the collection $\mathcal{O} := \{A^{-1}(x) : x \in X\}$ of open subsets of X forms a cover of X. By compactness, X can be covered by a finite subcollection $\mathcal{O}_f := \{A^{-1}(x_i) : i = 1, ..., n\}$. Let $\{\lambda_i : X \to [0,1]\}_{i=1}^n$ be a continuous partition of unity, subordinated to the finite cover $\mathcal{O}_f$ and define a continuous single-valued mapping $s : X \to X$ by:

$$s(x) := \sum_{i=1}^{n} \lambda_i(x)x_i, \text{ a convex combination, } \forall x \in X.$$

Since, for any given $x \in X$, $\lambda_i(x) \neq 0 \Rightarrow x \in A^{-1}(x_i) \Leftrightarrow x_i \in A(x) \Rightarrow s(x) \in A(x)$ because $A(x)$ is convex. The mapping s is finite dimensional, because $s(X) \subseteq C = Conv\{x_1, ...x_n\}$ a convex finite dimensional polytope $\subseteq X$. Now, the restriction/compression map $A_f := A|_C \cap C : C \to 2^C$ defined by:

$$A_f(x) = A(x) \cap C, \forall x \in C,$$

is certainly a $\mathbf{\Phi}^*-$map ($\forall x \in C,\ A_f^{-1}(x)$ is open in X, thus open in C, moreover $s(x) \in A_f(x) \neq \emptyset$ and $A_f(x)$ is convex). It is clear that a fixed point for A_f is also a fixed point for A. The Browder-Ky Fan fixed point theorem has thus been reduced to finite dimensional polytopes.

Second, Let $f = Id_C$ and $g : C \to C$ be a continuous selection of the map A_f. For a fixed but arbitrary $\epsilon > 0$, let $\Phi_\epsilon, \Psi_\epsilon : C \to 2^C$ be the two $\mathbf{\Phi}^*$ enlargements of f and g provided by Lemma 2.2.1, i.e.

$$f(x) \in \Phi_\epsilon(x) \subset B_\epsilon(f(B_\epsilon(x))), \forall x \in C, \tag{2.3}$$

$$g(y) \in \Psi_\epsilon(y) \subset B_\epsilon(g(B_\epsilon(y))), \forall y \in C. \tag{2.4}$$

By Theorem 2.2.8, for $\Phi_\epsilon, \Psi_\epsilon^{-1}$ there exists $(x_\epsilon, y_\epsilon) \in C \times C$ with:

$$y_\epsilon \in \Phi_\epsilon(x_\epsilon) \cap \Psi_\epsilon^{-1}(x_\epsilon).$$

By compactness, as $\epsilon \to 0$, a subset of (x_ϵ, y_ϵ) converges to some (x_0, y_0) in $C \times C$. Since f and g are continuous, inclusions 2.3 and 2.4 imply that $x_0 = g(y_0) \in A^f(y_0)$ and $y_0 = f(x_0) = x_0$, so $x_0 \in A_f(x_0)$. $\square$

We proceed now to close the lower loop of Figure 1.1 by showing that that Theorem 2.2.9 implies the KKM principle of Ky Fan (implication (8)).

Assume $\bigcap\limits_{x \in X} \Gamma(x) = \emptyset$, i.e $Y = \bigcup\limits_{x \in X} (Y \setminus \Gamma(x))$. Define $\tilde{A}, A : Y \to 2^Y$ by:

$$\tilde{A}(y) = X \setminus \Gamma^{-1}(y),$$

and

$$A(y) = Conv(X \setminus \Gamma^{-1}(y)).$$

Clearly $\tilde{A} \subseteq A$. Note $\forall y \in Y, \exists x \in X$ such that $y \in Y \setminus \Gamma(x) \Leftrightarrow y \notin \Gamma(x) \Leftrightarrow x \notin \Gamma^{-1}(y) \Leftrightarrow x \in X \setminus \Gamma^{-1}(y)$ i.e

$$\tilde{A}(y) \neq \emptyset \quad \text{for all} \quad y \in Y.$$

For each $z \in Y$ we have:

$$y \in \tilde{A}^{-1}z \Leftrightarrow z \in \tilde{A}y \Leftrightarrow z \in X \backslash \Gamma^{-1}y \Leftrightarrow z \notin \Gamma^{-1}y \Leftrightarrow y \notin \Gamma z,$$

if $z \in X$ then $\tilde{A}^{-1}z = Y \backslash \Gamma z$ which is open. If $z \in Y \backslash X$ then $\tilde{A}^{-1}z = \emptyset$ which is open too. So $\tilde{A}^{-1}(y)$ is an open set in Y. This implies that $A \in \Phi^*(Y)$. By Theorem 2.2.9, there exists y_0 such that $y_0 \in A(y_0)$ i.e

$$y_0 \in conv\{X \setminus \Gamma^{-1}(y_0)\}.$$

We have $y_0 \in Conv\{x_i\}_1^n$, $\{x_i\}_1^n \subset X$ and $x_i \notin \Gamma^{-1}(y_0)$ for all $i = 1, ..., n$ which means $y_0 \notin \Gamma(x_i)$ for all i. So:

$$Conv\{x_1, ..., x_n\} \not\subseteq \bigcup_{i=1}^{n} \Gamma(x_i),$$

which contradicts the hypothesis that Γ is a KKM map.

The von Neumann equilibria which are saddle points for payoff functions, are a particular case of Nash equilibria. The theorem on the existence of Nash equilibria is one of the fundamental results in game theory. A simple proof of the Nash theorem was provided by Ky-Fan based on the KKM theorem. It uses the result on nonempty intersection of family of sets with convex sections. Following the exposition of Granas [22]:

Given a certain product $X = \prod_{i=1}^{n} X_i$ of topological spaces, let $X^j = \prod_{i \neq j} X_i$ and let $p_i : X \to X_i, p^i : X \to X^i$ denote their projections; write $p_i(x) = x_i$ and $p^i(x) = x^i$. Given $x, y \in X$ we let

$$(y_i, x^i) = (x_1, ..., x_{i-1}, y_i, x_{i+1}, ..., x_n).$$

Theorem 2.2.10. ***(Geometrical result of Ky Fan)***

Let $X_1, X_2, ..., X_n$ be nonempty compact convex sets in linear topological spaces and let $A_1, A_2, ..., A_n$ be n subsets of X such that

(i)for each $x \in X$ and each $i = 1, 2, ..., n$,

$$A_i(x) = \{y \in X | (y_i, x^i) \in A_i\}$$

is convex and nonempty;

(ii) for each $y \in X$ and each $i = 1, 2, ..., n$

$$A^i(y) = \{x \in X | (y_i, x^i) \in A_i\}$$

is open. Then $\bigcap_{i=1}^{n} A_i \neq 0$

Proof. Define $G : X \to 2^X$ by $y \to X \backslash \bigcap_{i=1}^{n} A^i(y)$; one verifies that G is not a KKM-map and if a convex combination $w = \sum \lambda_i x_i \notin \bigcup G(x_i)$, then $w \in \bigcap_{i=1}^{n} A_i$. $\square$

As an immediate corollary:

Theorem 2.2.11. ***(The Nash Equilibrium Theorem)***

Let $X_1, X_2, ..., X_n$ be nonempty compact convex sets each in a topological vector space. Let $f_1, f_2, ..., f_n$ be n real-valued continuous functions defined on $X = \prod_{i=1}^{n} X_i$ such that for each $y \in X$ and each $i = 1, 2, ..., n$ the function $x_i \to$

$f_i(x_i, y^i)$ is quasiconcave on X_i. Then there is a point $y_0 \in X$ such that $f_i(y_0) = \max_{x_i} f_i(x_i, y_0^i)$.

Proof. Given $\epsilon > 0$, define for each $i = 1, 2, ..., n$

$$A_i^\epsilon = \{y \in X | f_i(y) > \max_{x_i \in X_i} f_i(x_i, y^i) - \epsilon\}.$$

Following from the quasiconcavity hypotheses on $f_i(x_i, y^i)$, A_i^ϵ is convex; and from their continuity, A_i^ϵ is open. Since, by Theorem 2.2.10 there exists, for each i and each ϵ a point $y_\epsilon \in \bigcap_{i=1}^{n} A_i^\epsilon$, and since X is compact, the net $y_\epsilon(\epsilon > 0)$ has a subnet converging to some point y_0 belonging to $\bigcap_{0<\epsilon<\epsilon_0} \bigcap_{i=1}^{n} A_i^\epsilon$ for some ϵ_0. This point y_0 satisfies the assertion of the theorem. □

With two agents with payoff functions $f_2 = -f_1$ we are in the situation of a zero-sum game and the Nash theorem reduces to the von Neumann theorem. Nash Theorem constitutes the theoretical basis for the second part of the thesis.

Deriving the KKM principle from the Berge-Klee theorem is not known yet. Horvath and Lassonde [24] used a topological version of the theorem of Berge-Klee to formulate a topological KKM principle where convexity is replaced by n-connectedness.

Definition 2.2.1. *If for a given positive integer n, Δ^n denotes the standard n-simplex whose vertices $\{e_0, ..., e_n\}$ form a canonical basis for $\mathbb{R}^{n+1}$ and $\partial\Delta^n$ is the boundary of Δ^n, a space X is* n-connected *if and only if every continuous function $f : \partial\Delta^n \to X$, extends continuously to a function $\tilde{f} : \Delta^n \to X$.*

The "topological Berge-Klee" reads as follow:

Theorem 2.2.12. ***(Topological Klee)***
A family of n closed convex sets in a topological vector space has a nonempty intersection if and only if the union of the n sets is $(n-2)$-connected and the intersection of every $n-1$ of them is nonempty.

It was shown in [24] that this result yields the equivalent formulation of the Brouwer fixed point theorem: the n-sphere S^n is not n-connected. Indeed, the n-dimensional faces of the n+1-simplex Δ^{n+1} form a family of $n+2$ closed convex sets in $\mathbb{R}^{n+2}$. Moreover, every intersection of $n+1$ of them is nonempty, but the whole intersection is empty. Hence, their union, which is $\partial\Delta^{n+1}$ is not n-connected. Since $\partial\Delta^{n+1}$ is homomorphism to S^n, S^n is not n-connected either.

which is equivalent to KKM principle. Note that since every convex set in a topological vector space is contractible, hence n-connected, the topological klee theorem implies the Berge-Klee theorem.

We show that the Berge-Klee intersection theorem is equivalent to a weaker version of the KKM principle, namely a convex KKM principle. This was pointed at by Granas and Lassonde [23] in the context of super-reflexive Banach space (they called this weaker KKM principle "elementary"). Their proof is based on a result on the minimization of quasiconvex, coercive, and l.s.c. functionals on closed convex subsets of super-reflexive spaces. We formulate below this result for arbitrary topological vector spaces using an extension of Berge-Klee's result to arbitrary topological vector space due to Ghouila-Houri [20].

Theorem 2.2.13. ***(Convex KKM principle)***

Let $\Gamma : X \to 2^E$ be a set-valued map defined on an arbitrary subset X of a real topological vector space E, with Γ be a KKM map and $\Gamma(x)$ closed convex for all $x \in X$. Then $\{\Gamma(x)\}_{x\in X}$ has finite intersection property.

We show first that this weaker version of the KKM principle implies the Berge-Klee theorem following an argument of S. Park [37].

Assume C and $C_1, ..., C_n$ are closed convex sets in a Euclidean space satisfying, $C \cap \bigcap_{i=1, i\neq j}^{n} C_i \neq \emptyset$ (for $j = 1, 2, ..., n$), and $C \subseteq \bigcup_{i=1}^{n} C_i$.

As $C \cap \bigcap_{i=1, i\neq j}^{n} C_i \neq \emptyset$ there exists $x_j \in [C \cap \bigcap_{i=1, i\neq j}^{n} C_i]$ for $j = 1, 2, ..., n$. Put $D = \{x_j\}_{j=1}^{n}$, $D \subset C$ and as C is convex, $Conv\{x_j\}_{j=1}^{n} \subset C$. Also for all $j \neq i$,

$x_j \in C_i$, which implies $Conv\{x_j\}_{j=1,j\neq i}^{n} \subset C_i$. Put $A_i = Conv\{x_j\}_{j=1,j\neq i}^{n}$ for $i = 1, 2, ..., n$. Define $\Gamma : D \to 2^C$ with $\Gamma(x_i) = C_i \cap C$ for $i = 1, 2, ..., n$.

$\Gamma(x_i)$ is closed in C for all $x_i \in D$. Also $Conv(D) \subset C \subset \bigcup\limits_{i=1}^{n} C_i \cap C = \bigcup\limits_{i=1}^{n} \Gamma(x_i)$. More over for each $\{x_{i_1}, ..., x_{i_k}\} \subset D$ we have $Conv\{x_{i_1}, ..., x_{i_k}\} \subset A_{i_j} \subset C_{i_j} \cap C = \Gamma(x_{i_j})$ for some $j \neq 1...k$. Hence $Conv\{x_{i_1}, ..., x_{i_k}\} \subset \bigcup\limits_{i=1}^{n} \Gamma(x_i)$, which shows that Γ is a KKM map.

By Theorem 2.2.6, $\bigcap\limits_{i=1}^{n} \Gamma(x_i) \neq \emptyset$, which implies that $\bigcap\limits_{i=1}^{n} C_i \neq \emptyset$. Which is the conclusion of the Berge-Klee theorem.

We end by showing that, conversely, the Berge-Klee theorem implies the convex KKM principle (implication 11 in Figure 1.1). Let us mention first that the Berge-Klee theorem holds in arbitrary topological vector spaces. More precisely, we have:

Theorem 2.2.14. *(Ghouila-Houri [20])*

Let $C_1, ..., C_n$ be closed convex sets in a topological vector space satisfying:

(i) each k of them $1 \leq k < n$ have a common point;

(ii) $\bigcup\limits_{i=1}^{n} C_i$ is convex;

Then $\bigcap\limits_{i=1}^{n} C_i \neq \emptyset$.

The proof is quite straightforward: let $x_i \in \bigcap\limits_{j\neq i} C_j$ and apply the theorem of Berge-Klee to the finite dimensional sets $C = Conv\{x_i\}, \tilde{C}_i = C \cap C_i$.

Now, using Theorem 2.2.14 we derive the convex KKM principle. Assume $\Gamma : X \to 2^E$ is a KKM-map with $\Gamma(x)$ closed convex. We show by induction on n that $Conv\{x_1, ..., x_n\} \cap \bigcap\limits_{i=1}^{n} \Gamma(x_i) \neq \emptyset$.

Step 1: $x_1 = Conv\{x_1\} \subset \Gamma(x_1)$ since Γ is a KKM-map.

Step 2: Assume the conclusion holds for any set of $n = k$ elements.

Step 3: Let $n = k+1$. Put $c = Conv\{x_1, ..., x_n\}$ and $C_i = \Gamma(x_i) \cap C$ for $i = 1, ..., n$. Since Γ is a KKM-map $C \subset \bigcup\limits_{i=1}^{n} \Gamma(x_i) \Rightarrow C = \bigcup\limits_{i=1}^{n} (C \cap \Gamma(x_i)) = \bigcup\limits_{i=1}^{n} C_i$. C is convex so $\bigcup\limits_{i=1}^{n} C_i$ is convex, and by step 2 of induction we have $Conv\{x_1, ..., \hat{x}_i, x_n\} \cap$

$\bigcap\limits_{j=1, j\neq i}^{n} \Gamma(x_j) \neq \emptyset$. By Theorem 2.2.14, $\bigcap\limits_{i=1}^{n} (C \cap \Gamma(x_i)) \neq \emptyset$ which implies that $\bigcap\limits_{i=1}^{n} \Gamma(x_i) \neq \emptyset$.

Chapter 3

Sponsored Search

3.1 Introduction to sponsored search

Web search is a significant technology for navigating the Internet. It has become a fundamental part of the online experience of Internet users which is provided for free. There were 10.27 billion web searches conducted in January 2010 in the U.S. Sponsored search is the delivery of relevance with advertisements as part of the search experience. Sponsored search is conducted through search engines, such as Google, Yahoo! and MSN. In Figure 3.1 you can see a layout of Google result after searching for keyword "safety gloves". In the main part of the result page, there are a list of links to the web pages which are relevant to the keyword "safety gloves". There are also three links on top of the page, and six links on right side of the page, which are the advertisements. Search engines earn money by displaying advertisements alongside the result page of a user inquiry.

Sponsored search satisfies users' desire for relevant search results and advertisers' need for increasing traffic to their web sites, and it is now considered to be among the most effective marketing tools available. Sponsored search has become a big business among different kinds of advertising. Google generated roughly $22 billion in revenue in 2009 that is almost 97% of its revenue.

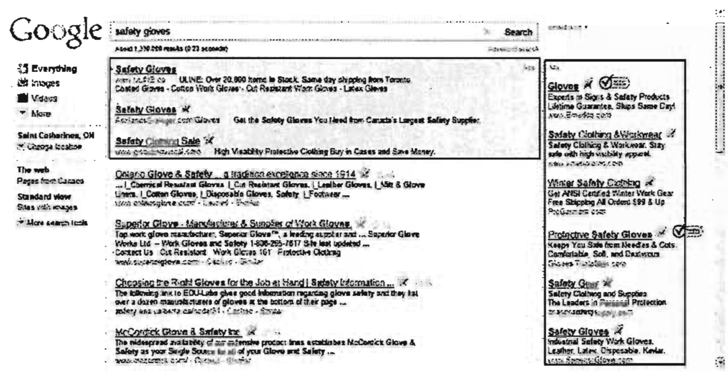

Figure 3.1: Sample example of Google result page for query "safety gloves"

Search engines provide spaces for advertisements. For example Google provides advertisers up to three spaces above the result page, and up to eight spaces besides the result page. We call each of these spaces a *slot*. Advertisements consist of a title, a text description, and a hyperlink to the advertiser's web page. Advertisers like users to click on their link and visit their web page. In this case they have a chance to introduce their products to users, and hopefully convince them to make a purchase. Usually advertisements get more attentions and more clicks when they appear in higher slots. As a result, advertisers generally prefer to be shown in higher slots rather than lower ones.

Every day hundreds of thousands of advertisers compete for slots alongside several million of search queries. The number of advertisers is usually more than the number of slots available. Search engines need to pick some of the advertisements to display them on the result page, so they conduct an auction among advertisers to allocate slots. Each advertiser specifies: (i) a list of keywords she is interested to advertise for them; (ii) a bid, which is the maximum money she is willing to pay to the search engine for each advertisement; (iii) a total maximum daily or weekly budget. Every time a user searches for a keyword, an auction takes place among the advertisers who are interested in that keyword and have not exhausted their budgets. A search engine scores advertisers based on their private scoring scheme, and then allocates slots in decreasing order of scores, so that the advertiser with the highest score is shown in the first slot, and so on.

There are different kinds of charging schemes. In the most popular one, advertisers pay the search engine only when a user clicks on their advertisement, and do not pay if their advertisement is displayed but not clicked on. We call this *pay per click auction*. In other auctions, advertisers *pay per impression* or *per purchase* (by impression we mean placement on the screen).

3.2 Generalized Second Price (GSP) auctions

As we mentioned in the previous section, every time a user searches for a keyword, an auction takes place among the advertisers who are interested in that keyword. Let A be the number of advertisers and S be the number of available slots. Assume that $x_{s,a}$ is the probability that a user clicks on the s'th slot when it is occupied by advertiser a. We call $x_{s,a}$ the *click through rate* (CTR) for advertiser a in slot s. As we explained, advertisers prefer to be shown in the first slot rather than second slot and so on. Therefore, we have $x_{s,a} \geq x_{s+1,a}$ for $s = 1, ..., S-1$. Notice that this inequality holds on the CTRs of the same advertiser. According to the quality of an advertiser, the search engine assigns a weight w_a to each advertiser a. This weight is based on the probability of an advertiser being clicked on by users. If b_a is the bid of advertiser a, then search engine allocates a score $\mu_a = w_a b_a$ to advertiser a and orders all advertisers according to their scores in decreasing way. Then the advertiser with the highest score is ranked one, and the advertiser with second highest score is ranked two and so on. For simplicity, we renumber advertisers so that advertiser s obtains slot s. The payment of advertiser s denoted by p_s is equal to the bid of advertiser $s+1$ i.e we have $p_s = b_{s+1}$.

We call this mechanism *generalized second price* (GSP). If $w_a = 1$ for every a, then the auction is called *rank by bid*, and if $w_a = x_{1,a}$, the auction is called *rank by revenue* [36].

3.3 Externality in sponsored search auctions

In GSP auctions we assign a CTR to each advertiser. This CTR depends on the quality of the advertiser and the quality of the slot she occupies. Although this assumption is used by most studies on sponsored search auctions, it is not realistic. Indeed, the CTR depends on the identity and position of the other advertisers who are displayed on the result page as well. In fact, advertisers have externality effects on each other, which influence their CTRs and consequently affect their bidding behavior.

To study the externality effects, we look at these auctions from the perspective of the users. Generally users do not click on all the advertisements that appear on the result page. They click on the advertisements that look good enough to satisfy the reason of their inquiry. When a user clicks on an advertisement, she gathers some information toward her search. This information may compensate the user's needs, and thus cause her not to continue looking at other advertisements. Also the user may get tired of the search, if the advertisements she has read appear to be poorly related to the search term. That would lead the user not to look at the rest of the advertisements.

These externality effects were shown in the experimental work of R. Gomes and N. Immorlica [21] on on-line auctions. They gathered the data consisting of impression and clicking records associated to queries on the keywords "iPods", "diet pills" and "Avg antivirus" in Microsoft's Live Search. These data are related to the inquiries that happened between August 1st and November 1st of 2007 (Table 3.1).

As you can see in Table 3.2, these three advertisers appear in almost all three slots available for advertising.

In Table 3.3, $F_A = F_A(\{\emptyset\})$ is the probability that a user clicks on the link of advertiser A, when she did not click on any other advertisements before; $F_A(\{B\})$ is the probability that a user clicks on the link of advertiser A, when she has already clicked on advertiser B's link. $F_A(\{B, C\})$ is the probability that a user clicks on the link of advertiser A, when she has already clicked on the links of advertisers B and C.

keyword	advertisers	# of observations
iPod	(A):store.apple.com (B):cellphoneshop.net (C):nextag.com	8,398
diet pill	(A):pricesexposed.net (B):dietpillvalueguide.com (C):certiphene.com	4,652
Avg antivirus	(A):Avg-Hq.com (B):avg-for-free.com (C):free-avg-download.com	1,336

Table 3.1: Keywords and advertisers. For the keyword iPod, there are three advertisers competing together. (A):Apple Store (www.store.apple.com); (B):the online retailer of electronics Cell Phone Shop (www.cellphoneshop.net); (C) the price research website Nextag (www.nextag.com). In this sample, there were 8398 observations for keyword iPod.

slot	iPod	diet pill	antivirus
first	(A): 6,460 (76.92%) (B): 1,864 (22.20%) (C): 74 (0.88%)	(A): 1,912 (41.10%) (B): 908 (19.52%) (C): 1,832 (39.38%)	(A): 1,233 (92.29%) (B): 71 (5.31%) (C): 32 (2.40%)
second	(A): 1,438 (17.12%) (B): 5,826 (69.37%) (C): 1,134 (13.50%)	(A): 1,848 (39.72%) (B): 1,988 (42.73%) (C): 816 (17.54%)	(A): 88 (6.59%) (B): 674 (50.45%) (C): 574 (42.96%)
third	(A): 26 (0.31%) (B): 22 (0.26%) (other): 7,400 (88.12%)	(A): 472 (10.15%) (B): 692 (14.88%) (other): 2,820 (60.62%)	(A): 9 (0.67%) (B): 21 (1.57%) (other): 951 (71.18%)

Table 3.2: Distribution of Advertisers per slot.

We notice that $F_B = 0.08 > 0.04 = F_B(\{A\})$ and $F_C = 0.10 > 0.04 = F_C(\{A\})$. It shows that a random user is less willing to click on Cell phone shop or Nextag links when she has already clicked on Apple store. This implies that the information provided by Apple store for users, reduce their appeal to click on other advertisements.

3.4 Accommodating externality requests

We saw that in reality the CTRs depend on the identity and position of the other advertisers who are displayed on the result page. It makes sense that advertisers like

keyword	iPod	diet pill	antivirus
F_A	0.210	0.210	0.151
$F_A(\{B\})$	0.250	0.232	0.00
$F_A(\{C\})$	-	0.317	-
$F_A(\{B,C\})$	-	0.664	-
F_B	0.087	0.150	0.206
$F_B(\{A\})$	0.030	0.146	0.364
$F_B(\{C\})$	-	0.663	-
$F_B(\{A,C\})$	-	0.334	-
F_C	0.104	0.051	0.215
$F_C(\{A\})$	0.040	0.052	0.242
$F_C(\{B\})$	0.095	0.088	0.121
$F_C(\{A,B\})$	0.327	0.664	0.125

Table 3.3: Estimate of the ordered search model.

to bid based on other advertisements on the result page. However no search engine accommodates such requests. Below we see why adding externality to one of the most basic models would cause extreme complication to the mechanism.

First we present a model that was introduced by Hal R.Varian [40], in which $w_a = 1$ for all $a \in A$, and CTRs only depend on the location of the slots (and not on the advertisers). Let x_s denotes the CTR assigned to slot s which is independent of the advertiser who is displayed in this slot.

Recall that A and S denote the number of advertisers and slots, respectively. As we said before, generally a slot with higher position has a higher CTR. So assume that, $x_1 > x_2 > \ldots > x_S$, and $x_s = 0$ for all $s > S$. Each advertiser is assigned a value $v_a > 0$ for $a = 1, ..., A$ which is the expected profit for advertiser a if it has been clicked on by a user. Therefore the expected profit for advertiser a when it is displayed in slot s is $u_{as} = v_a x_s$.

Recall that b_a is the bid of advertiser a, and as $w_a = 1$ the score of advertiser a is $\mu_a = b_a$. So the first slot will be given to the advertiser with highest bid and the second slot to the advertiser with second highest bid, and so on. Like before we renumber advertisers to have $a = s$, for all $a = 1, ..., S$. The payment of advertiser s is $p_s = b_{s+1}$. Therefore, the utility of advertiser s is $(v_s - p_s)x_s$.

Now we define the *Nash equilibrium* and *symmetric Nash equilibrium* concepts for this model. We say that there is a Nash equilibrium for a game when none of the players can improve their payoff by changing their strategy when other players stick to their past strategies. We will present the following formal definitions of Nash equilibrium and symmetric Nash equilibrium for this model. Notice that in this model, the set of prices determines the bidding strategies of advertisers.

Definition 3.4.1. *A Nash equilibrium set of prices (NE) satisfies*

$$(v_s - p_s)x_s \geq (v_s - p_t)x_t \qquad for \quad t > s \tag{3.1a}$$

$$(v_s - p_s)x_s \geq (v_s - p_{t-1})x_t \qquad for \quad t < s \tag{3.1b}$$

where $p_t = b_{t+1}$

Inequality (3.1a) shows that none of the advertisers can improve their utility by moving their advertisements to a lower position and aiming for a lower cost. Inequality (3.1b) shows that they cannot improve their utility by moving their advertisements to a higher position and aiming for a higher CTR.

Definition 3.4.2. *A symmetric Nash equilibrium set of prices (SNE) satisfies*

$$(v_s - p_s)x_s \geq (v_s - p_t)x_t \qquad for\ all \quad t \quad and \quad s \tag{3.2}$$

Table 3.4 shows an example of an auction with a SNE. In this example we have 6 advertisers and 4 slots. None of the advertisers can increase their utility by changing their bid. For instance if we put $s = 2$ and $t = 1$ in Inequality 3.2, then $40 = (15 - 7)5 = (v_2 - p_2)x_2 \geq (v_2 - p_1)x_1 = (15 - 14)12 = 12$. Hence if she increases her bid in order to move up to the first slot then her utility will drop 28 units. If she decides to move her advertisements down to a lower slot, let's say third slot, her utility will be $(15 - 3)3 = 36$ which is less than 40.

The definition of SNE derives some properties for this model:

Property of non-negative surplus: *In an SNE,* $v_s \geq p_s$.

slot	CTR	value	bid	payment
1	12	20	16	14
2	5	15	14	7
3	3	10	7	3
4	2	5	3	1
5	0	1	1	0
6	0	0.5	0.25	0

Table 3.4: Example of an auction with SNE.

Proof. In inequality 3.2, put $t = S + 1$. Then we have $(v_s - p_s)x_s \geq (v_s - p_{S+1})x_{S+1}$. As $x_{S+1} = 0$, thus $v_s - p_s \geq 0$, which implies $v_s \geq p_s$. □

Property of monotone values: *In an SNE,* $v_{s-1} \geq v_s$.

Proof. By the definition of SNE for s and t we have $(v_s - p_s)x_s \geq (v_s - p_t)x_t$, which is equivalent to $v_s(x_s - x_t) \geq p_s x_s - p_t x_t$. By rewriting this inequality for t and s we will have $v_t(x_t - x_s) \geq p_t x_t - p_s x_s$. Adding these two inequalities gives us $(v_t - v_s)(x_t - x_s) \geq 0$, which shows that v_t and x_t should be in the same order. □

Property of monotone prices: *In an SNE, if* $v_s > p_s$ *then* $p_{s-1} > p_s$.

Proof. We know that $p_{s-1} \geq p_s$, . We want to show that this inequality is strict when we have $v_s > p_s$. In inequality 3.2, put $t = s - 1$, then $p_{s-1}x_{s-1} \geq p_s x_s + v_s(x_{s-1} - x_s)$. Since by non-negative surplus property $v_s > p_s$, hence $p_{s-1}x_{s-1} > p_s x_s + p_s(x_{s-1} - x_s) = p_s x_{s-1}$. By crossing x_{s-1} from both sides of this inequality we have $p_{s-1} > p_s$. □

Property of inclusion: *SNE* $\subset$ *NE.*

Proof. The first inequality of NE is the same as SNE so we just need to show that we have the second inequality as well. According to SNE for all s and t we have $(v_s - p_s)x_s \geq (v_s - p_t)x_t$. Since $p_{t-1} \geq p_t$ we will have $(v_s - p_s)x_s \geq (v_s - p_{t-1})x_t$. □

Property of one step solution: *If a set of prices satisfies the SNE inequalities for $s-1$ and s, then it satisfies these inequalities for $s-1$ and $s+1$.*

Proof. Put s and $s+1$ in inequality 3.2, then $v_s(x_s - x_{s+1}) \geq P_s x_s - p_{s+1}x_{s+1}$. Since $v_{s-1} \geq v_s$, then $v_{s-1}(x_s - x_{s+1}) \geq P_s x_s - p_{s+1}x_{s+1}$. Adding this inequality to $v_{s-1}(x_{s-1} - x_s) \geq P_{s-1}x_{s-1} - p_s x_s$ (it comes from putting $s-1$ and s in inequality 3.2), we get $v_{s-1}(x_{s-1} - x_{s+1}) \geq P_{s-1}x_{s-1} - p_{s+1}x_{s+1}$. This shows that we have SNE for $s-1$ and $s+1$. □

Now we extend Hal R. Varian model by adding a parameter to this model. This in a sense, is the most basic accommodation of externality request by users. We will define the new model and then we will study the properties of this model.

Consider the auction model we presented before. Assume that among all advertisers there is one advertiser a which is in competition with agent $\hat{a}$. Advertiser a does not want to be shown after her competitor $\hat{a}$ on the result page, i.e. if it happens that a falls after $\hat{a}$ she prefers not to be shown in the page. Now we present the modified definition of NE and SNE for this extended model.

Definition 3.4.3. *A Nash equilibrium set of prices (NE) satisfies*

$$(v_s - p_s)x_s \geq (v_s - p_t)x_t \quad \text{for} \quad t > s \quad \text{except when } s = a < \hat{a} \leq S \text{ and } \hat{a} \leq t \leq S \tag{3.3}$$

$$(v_s - p_s)x_s \geq (v_s - p_{t-1})x_t \quad \text{for} \quad t < s \tag{3.4}$$

where $p_t = b_{t+1}$

Definition 3.4.4. *A symmetric Nash equilibrium set of prices (SNE) satisfies*

$$(v_s - p_s)x_s \geq (v_s - p_t)x_t \text{ for all } t \text{ and } s \text{ except when } s = a < \hat{a} \leq S \text{ and } \hat{a} \leq t \leq S$$

The SNE of this modified model differs from the SNE of Hal R. Varian model only when $s = a < \hat{a} \leq S$ and $\hat{a} < t \leq S$. In this case, advertiser a may improve her utility by moving to a slot below $\hat{a}$, but she does not move because of her preference not to be shown after her competitor. Here we discuss the properties of this new model according to Definition 3.4.4:

Property of non-negative surplus holds. Since when $t > S$ the SNE of this modified model is the same as the SNE of Hal R. Varian model, by putting $t = S+1$, like before we can show that this model satisfies this property.

Property of monotone prices holds. Since when $t < s$ the SNE of this modified model is the same as the SNE of Hal R. Varian model, by putting $t = s-1$, the same as before we can show that this model satisfies this property. Clearly, SNE $\subset$ NE for this model, so the property of inclusion is satisfied by this model too.

Other two properties do not hold any more. To see this, we present an example in Table 3.5. In this example we have 6 advertisers and 4 slots, and a is the second advertiser and $\hat{a}$ is the third one. As you can see none of the advertisers can improve their utilities by changing their bid, except for the case that advertiser a wants to move down to slot 3 or 4. Therefore we have SNE for this model. In this example $v_2 < v_3$, so we do not have monotone value property.

slot	CTR	value	bid	payment
1	27	20	17	16
2	13	16	16	14
3	10	19	14	10
4	5	10	10	5
5	0	5	5	0
6	0	4	4	0

Table 3.5: Example of an auction with SNE.

The one step property does not hold, consider the case in which $s - 1 = a$ and $S \geq s + 1 = \hat{a}$.

Chapter 4

Network of on-line advertisers

As we showed in Section 3.4, accommodating externality seems to add complexities to the model. Moreover, search click data due to the privacy issues is extremely difficult to collect and search engines are reluctant to disclose such statistical data. So in order to have a better understanding of the interaction among advertisers, we study the underlying graph structure of the Google advertisements network. This graph is dynamic and consists of thousands of vertices. Thus, to study its structure, we apply the existing models to large-scale networks.

4.1 Large-scale networks

The study of large-scale networks falls in the field of graph theory. We represent a network by a graph. In this chapter we use the definitions and notations from [45]. A *graph* G is a triple consisting of a *vertex set* $V(G)$, an *edge set* $E(G)$, and a relation that associates with each edge two vertices, called its *endpoints* (see Figure 4.1 for an example). A graph G is a *directed graph* if there is a function assigning each edge an ordered pair of vertices. The first vertex of the ordered pair is the *tail* of the edge, and the second is the *head*; together, they are the endpoints. We say that an edge is an edge *from* its tail *to* its head.

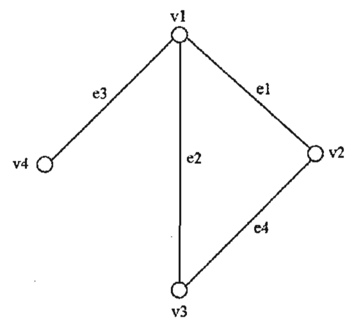

Figure 4.1: A graph with 4 nodes and 4 edges, $V = \{v_1, ..., v_4\}$, $E = \{e_1, ..., e_4\}$. The assignment of endpoints to edges can be read from the picture.

Examples of studied large-scale networks are:

- *Science collaboration network* whose vertices are scientists and an edge exists between two scientists if they have written an article together.
- *World Wide Web* whose vertices are web pages and the edges represent the hyperlinks (URLs) that point from one web page to anotheother.
- *The web of human sexual contacts* is a network of people who are connected to each other if they had a sexual relationship.
- *On-line Social Networks (OSNs)* such as Facebook, whose vertices are users of this service and there is an edge between two users if they are friends.

Gathering information about large-scale networks used to require a vast type of resources. Recently, some developments have caused dramatic improvements in studying large-scale networks. The computerization of data gathering enables us to analyze enormous database of these networks, and also the improvement of computational power provides researchers with the ability to investigate on large-scale networks.

Following these developments, studying the structural properties of a variety of real-world networks became possible. Many new concepts and measures have been proposed and studied on such large-scale networks. Among these concepts *small-world* has been the main focus of most studies.

Before we define the small-world networks, we present some preliminary definitions. In a graph $G(V, E)$:

- When u and v are the endpoints of an edge, they are *adjacent* and are *neighbors*. For example in Figure 4.1, v_1 and v_2 are the neighbors of vertex v_3.

- The *degree* of vertex v, is the number of edges for which v is an endpoint. The degree of vertex v is denoted by d_v. For example in Figure 4.1, $d_{v_1} = 3$. In directed graphs we define *in-degree* and *out-degree* for each vertex v as well. The in-degree of a vertex v is the number of edges for which v is the head and is denoted by d_{in_v}. The out-degree of vertex v is the number of edges for which v is the tail and is denoted by d_{out_v}.

- If v_0 and v_k are two vertices, then a *path* between them is a list $v_0, e_1, v_1, ...,$ v_{k-1}, e_k, v_k of vertices and edges such that, for $1 \leq i \leq k$, the edge e_i has endpoints v_{i-1} and v_i. The length of a path is the number of edges it contains. For example in Figure 4.1, $v_4, e_3, v_1, e_1, v_2, e_4, v_3$ is a path between vertices v_4 and v_3 of length 3. In a directed graph $v_0, e_1, v_1, ..., v_{k-1}, e_k, v_k$ is a path between v_0 and v_k if and only if v_{i-1} and v_i are head and tail of edge e_i, respectively, for all $1 \leq i \leq k$.

- The *distance* between two vertices is the length of the shortest path between these two vertices. For example in Figure 4.1, the distance between v_4 and v_3 is 2.

- A *clique* in a graph is a set of pairwise adjacent vertices. A clique with n vertices has $\binom{n}{2} = n(n-1)/2$ edges. In directed graphs, a clique with n

vertices has $2\binom{n}{2} = n(n-1)$ edges.

- In undirected graphs, the *clustering coefficient* of vertex v with degree d_v is defined as:

$$C_v = \frac{E_v}{\binom{d_v}{2}},$$

and for directed graphs as:

$$C_v = \frac{E_v}{2\binom{d_v}{2}}.$$

where E_v is the number of edges that exist among neighbors of vertex v, and d_v is the degree of v. The clustering coefficient of a graph is the average of the clustering coefficients of all its vertices. For example in Figure 4.1, $c_{v_1} = 0.33$ and $c_{v_2} = 1$.

Definition 4.1.1. *[5] A small-world network is a network with the following properties: (i) its average distance increases logarithmically with the number of vertices; and (ii) it has high clustering coefficients.*

The first property describes that, despite the large size of the small-world networks, there is a short path between most pairs of vertices. In simple terms, the average distance of these networks is small. This property was first studied by the social psychologist Stanley Milgram (1967). He described the *six degrees of separation* concept, which says that there is a path of length of at most six between most pairs of people in the United States [32].

The second property of small-world networks is a measure of *cliquishness* of neighborhoods in these networks. It shows that these networks contain sub-networks which have connections between most pairs of vertices. This property was first uncovered by Wassermann and Faust (1994) under the name *fraction of transitive triples* [44]. When we say the clustering coefficient in these networks is high, we mean that the clustering coefficient is larger than the clustering coefficient of a random graph with the same number of vertices and edges.

There are three classes of small-world networks: (a) scale-free networks; (b) broad-scale networks; and (c) single-scale networks. We focus on small-world networks which are scale-free. Scale-free networks are networks whose degree distribution has the *power-law*.

The *degree distribution* $P(k)$ of a graph is the distribution of its degrees over the whole graph. We say that a graph G has power-law degree distribution if for every k:

$$P(k) \propto L(k)k^{-\gamma}$$

where $\gamma > 1$ and is called the *exponent* of the power-law. Also $L(k)$ is a slowly varying function, which is any function that satisfies $\lim_{k\to\infty} L(tk)/L(k) = 1$ with t constant, usually $L(k)$ is considered to be a constant. This property of $L(k)$ follows directly from the requirement that $P(k)$ be asymptotically scale invariant; thus, the form of $L(k)$ only controls the shape and finite extent of the lower tail. For instance, if $L(k)$ is the constant function, then we have a power-law that holds for all values of k (see Figures 4.2,4.3).

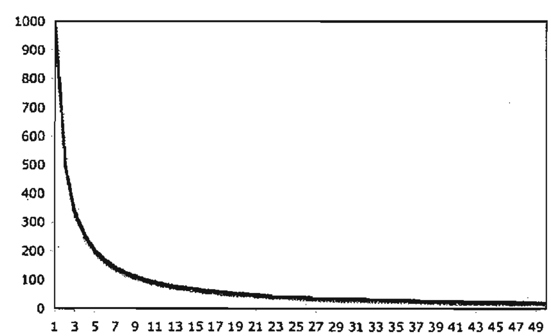

Figure 4.2: Sample of power-law degree distribution. X-axis is the degree and Y-axis is the number of vertices with that degree.

A highlight for scale-free networks is that they can be produced by a *preferential attachment* process (it is not the only process who produces scale-free networks [14]). A preferential attachment process is a process in which a new vertex is connected preferentially to the vertices with higher degrees. When we study the graph of scale-

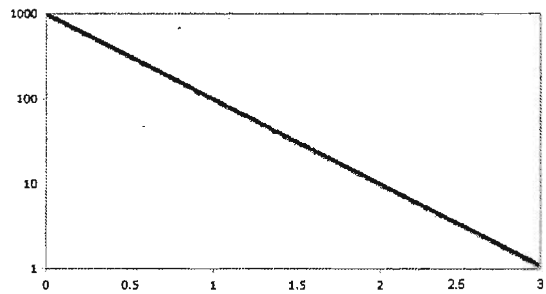

Figure 4.3: When power-law is plotted as a log-log curve then this gives a straight line. Such log-log plot is a standard way of testing whether any kind of behavior has a power-law distribution. X-axis is the degree and Y-axis is the LogLog number of vertices with that degree.

free networks, we observe few vertices with a high degree, with many vertices of small degrees connected to these high-degree vertices which we call *hubs*. If we randomly remove a vertex from the graph, as the number of small degree vertices are much higher than hubs, the probability that this vertex has a small degree is high. Therefore removing a random vertex will not change the shape of the graph. The phrase "scale-free" comes from this fact (see Figure 4.4).

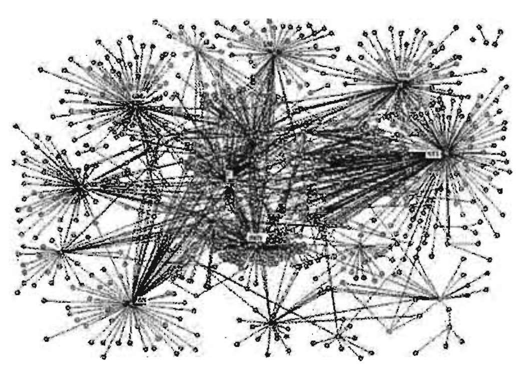

Figure 4.4: Example of a scale-free network.

Now we review some examples of large-scale networks which we introduced earlier. We will verify that they satisfy the property of small-world scale-free networks.

- **Science collaboration graph**

In the collaboration network, vertices are the scientists and an edge represents the existence of an article which has been written by two scientists together. It was Newman [34, 33] who studied physics articles between 1995 and 1999. He showed that this network is a small-world network. It has a small average distance of 4.0 and a very high clustering coefficient of 0.726. This network also falls in the scale-free class of small-world networks as it has power-law degree distribution with degree exponent of 1.2 (see Figure 4.5). A scientist who wants to write his/her first article in collaboration with other scientists, more likely will do it with more senior scientists who have written more articles in collaboration with others. This implies the preferential attachment property of this scale-free network.

Barabasi et al [8] did the same work of Newman on the mathematicians and neuroscientists articles between 1991 and 1998. They have average distance of 9.5 and 6, and clustering coefficients of 0.59 and 0.76, respectively. The degree distributions of both have power-law; the degree exponent of mathematicians is 2.1, and the degree exponent of neuroscientists is 2.5 (see Figure 4.5).

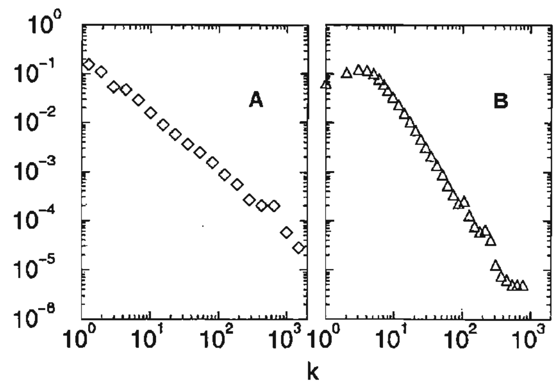

Figure 4.5: (A) Science collaboration network of physicists (B) Science collaboration network of neuroscientists. X-axis is the degree and Y-axis is the number of vertices with that degree.

- **World Wide Web**

  In the network of World Wide Web, vertices are web pages and the edges represent the hyperlinks that point from one web page to another. At the end of 1999, this network contained close to one billion vertices. This network is one of the most studied examples of small-world networks. Albert, Jeong and Barabási [3] showed that the average distance in a sample subgraph that consists of 325,729 vertices is 11.2. Since the edges of this network are directed, the network is characterized by both out-degree and in-degree distributions. In this sample, both degree distributions has power-law. The degree exponent of the out-degree distribution is 2.45, and the degree exponent of the in-degree distribution is 2.1. It was Adamic [1] who studied the clustering coefficient of this network. She made the graph of this network undirected by making each edge bidirectional. Her sample consisted of 153,127 vertices and had a clustering coefficient of 0.1078.

- **The web of human sexual contacts**

  Many sexually transmitted diseases, including AIDS, spread on a network of sexual relationships. Liljeros et al [30], focused on studying the network of sexual relations of 2,810 individuals, based on an extensive survey conducted in Sweden in 1996. In this network vertices are 2,810 Swedish that are connected to each other if they had a sexual relationship. Since the relationships does not last long, they analyzed the distribution of partners over a single year, obtaining for both females and males a power-law degree distribution with a degree exponent of 3.5 for female and 3.3 for males.

- **On-line Social networks (OSNs)**

  Facebook, Twitter, Orkut, MySpace and Cyword are examples of OSNs. The number of users of OSNs is almost half of all Internet's users [2]. The growth rate of these networks is considerably high. People use these networks to do social activities such as: finding friends with common interests; sharing pictures; discussing about their favorite subjects and so on.

Knowing about the structure of OSNs is interesting for sociologists and the owner of these services. A basic question is if the relations and the way that people grow their communities in these networks is similar to those in real world or not. The owner of these services earn money by letting companies advertise on their pages, and/or by offering their users to become a premium member through periodic payments (being a premium member lets the user to have access to more information about others). Therefore it is important for them to know how to improve their model, and how to design new applications in order to attract more users and consequently making more revenue. OSNs are another example of small-world networks.

Y. Ahn Studied the network of Cyworld [2]. Cyworld is a South Korean social network service. This OSN started its operation in September 2001, and had 12 million users by November 2005 (compare with population of South Korea which was 48 million at that time). Similar to other OSNs, Cyworld offers a space to its users to make a friend (called *ilchon*) online.

He reported that the average distance in Cyworld is 5.8 and the clustering coefficient is 0.16. The degree distribution of Cyworld has power-law as shown in Figure 4.6. This degree distribution consists of two parts, a rapid decaying $\gamma \sim 5$ region and a heavy tailed $\gamma \sim 2$ region (the division takes place between $k = 100$ and $k = 1000$).

4.1.1 A model for scale-free networks

The results discussed above demonstrate that many large-scale networks are scale-free. Here we are going to present a natural random process that makes a graph whose degree distribution has power-law.

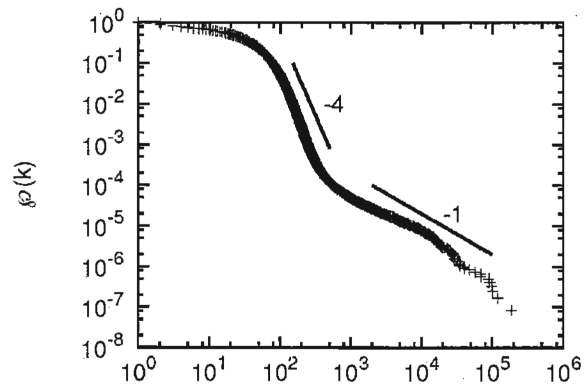

Figure 4.6: Degree distribution of Cyworld network. X-axis is the degree and Y-axis is the number of vertices with that degree.

There are different types of models that generate random graphs. One of these models is the model introduced by Barabási and Albert in 1999 [7]. There are two major differences between Barabási-Albert's model and the earlier models. First of all, earlier models do not change the number of vertices over time: they start making a random graph by randomly connecting a fixed number of vertices. In contrast, Barabási-Albert's model assumes that the number of vertices is not fixed. It starts making the random graph from a small number of vertices and then grows the graph by adding new vertices continuously. For example, Facebook grows exponentially in time by adding new users, and the science collaboration network constantly grows by the appearance of new researchers.

Secondly, earlier models assume that the probability that two vertices are connected to each other is independent of the degree of them. But as we mentioned before, scale-free networks exhibit preferential attachment. Barabási-Albert's model considers this fact by taking into account that connecting a new vertex to an existence vertex depends on the degree of that vertex. For example, a new web page is more likely to have hyperlinks to well-known web pages rather than less-known pages.

By taking the two facts of growth and preferential attachment into consideration, Barabási and Albert introduced their model, which led for the first time to a network with a power-law degree distribution.

The Barabási-Albert model

The process in Barabási-Albert model is as follows.

- Growth: Start with a small graph G_0 with no isolated vertices (m_0 is the number of vertices of G_0). At every time step, we add a new vertex with m ($< m_0$) edges that link this new vertex to m different vertices already present in the system.

- Preferential attachment: The probability $\Pi(i)$ that a new vertex will be connected to an existing vertex i depends on the degree of node i,

$$\Pi(i) = \frac{d_i}{\sum_j d_j}.$$

After t time steps, this procedure results in a network with $N = t + m_0$ vertices and mt edges. The degree distribution of this network has power-law with degree exponent of $\gamma \sim 3$ (see Figure 4.7). The degree exponent is independent of m the only parameter in the model.

4.2 A small-world network

As we said before, search engines such as Google, Yahoo!, Microsoft and AOL, earn money by selling advertising spaces in the result page of a user's inquiry. As you can see in Table 4.1, Google generated almost 70% of all revenue made by online advertising in 2009. So in order to study the network of sponsored search, we use the network induced by Google.

"Google advertisements network" is an example of a large-scale networks. The aim of this section is to analyze the structure of this network to show that it is a scale-free small-world network.

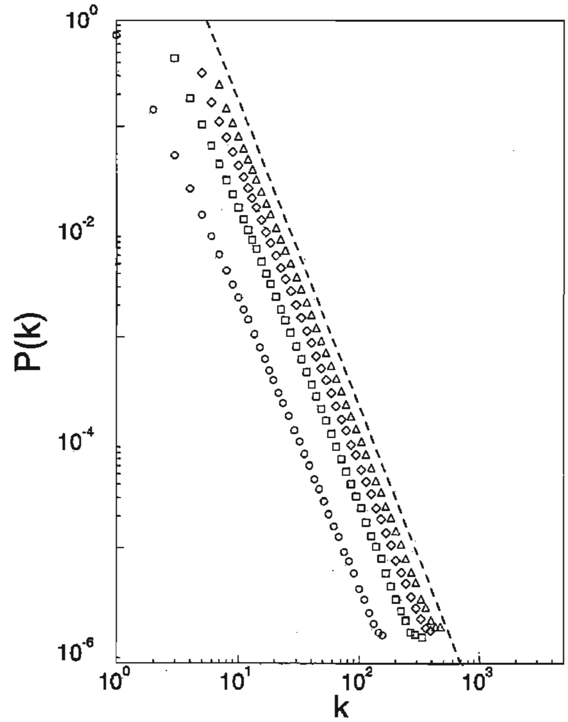

Figure 4.7: Degree distribution of the Barabási-Albert model, with $N = m_0 + t = 300{,}000$ and $\circ$, $m_0 = m = 1$; $\square$, $m_0 = m = 3$; $\diamond$,$m_0 = m = 5$; and $\triangle$, $m_0 = m = 7$. The slope of the dashed line is $\gamma = 2.9$, providing the best fit to the data. X-axis is the degree and Y-axis is the number of vertices with that degree.

Google	\$22,889
Yahoo	\$5,673
Microsoft	\$2,131
AOL	\$1,749
Total	\$32,442

Table 4.1: Online Advertising Revenues for 2009 (in millions)

We model this network by a directed graph. Google considers up to three slots above the result page (called *top-of-page* slots), and up to eight slots besides the result page (called *sidebar* slots) for advertisements. As you can see in Figure 3.1 each advertisement contains a URL address which is in green color. For example the first top-of-page advertisement is for ULINE (a shipping company) and it contains "www.ULINE.ca" which is the URL address of ULINE company. The vertices of this graph are these URLs displayed on the result page of inquiries. The advertisers are known by their URL addresses. Sometimes the URL address which is shown in the advertisement is the URL address of a specific department of the company. For example, look at the second top-of-page advertisement in Figure 3.1. This advertisement belongs to the AcklandsGrainger company which sales safety products. The URL address is shown in this advertisement is "AcklandsGrainger.com/Gloves" which is for gloves department of this company. In these cases, consider "AcklandsGrainger.com" as the name of the vertex, not "AcklandsGrainger.com/Gloves". The reason is that AcklandsGrainger company may advertise for both "safety gloves" and "safety boots" (or even more keywords), and the URL address that appears in each case is related to the relevant department. This consideration avoids duplicating advertisers.

There is an edge between two advertisers if they have been displayed on the result page of an inquiry together. We number slots such that the top slots are numbered top-down 1 to 3 and the 8 right-side slots are numbered 4 to 11. The edges are from advertisers in the slots with higher numbers to advertisers in the slots with lower numbers. For example in Figure 3.1 there are edges from "www.globalindustrial.com" to "www.ULINE.ca" and "AcklandsGrainger.com".

The list of vertices can be found by searching every possible keyword in Google. Google ADWords has a complete list of keywords that companies are interested to advertise for. In the Google ADWords data base there are 17 categories (such as apparel, beauty, computer,etc) of keywords and they all together contain 561,846 keywords.

We wrote a code in Python programming language. This code gets the list of the keywords as input (see Figure 4.8) and search each of them in Google. The output of this code is a list which contains the keywords followed by the URL addresses of the companies who advertised for these keywords (see Figure 4.9). After each keyword, there is a list of the URL addresses. These URLs are ordered according to their appearance in the result page of inquiry for that keyword. First the URLs of top-of-page advertisements and then the URLs of sidebar advertisements (both in top-down order). Because of the reasons that we mentioned before, we wrote this code in a way that it returns the URL addresses of the companies not the company's departments. Refer to Appendix (A) for the source code in Python.

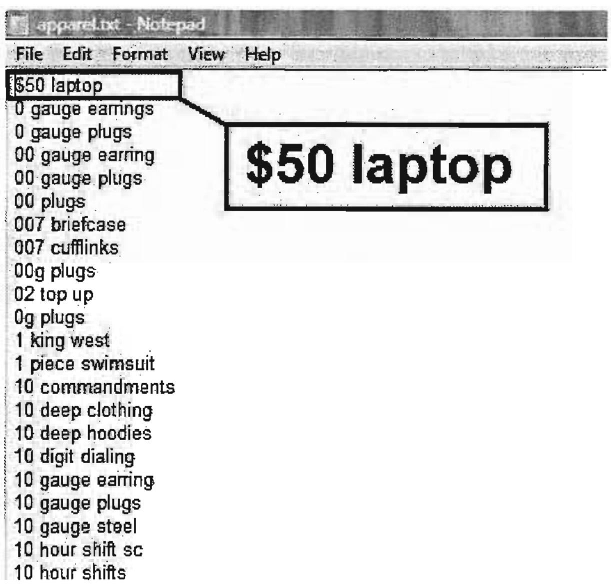

Figure 4.8: Sample input for Python code. It is a list of keywords that companies are interested to advertise for in Google.

In order to set up the vertex and edge sets, we wrote a C++ code (refer to Appendix (B) for the source code in C++). The input of this code is the output of the Python code. It implements the following algorithm.

For each keyword in the list:

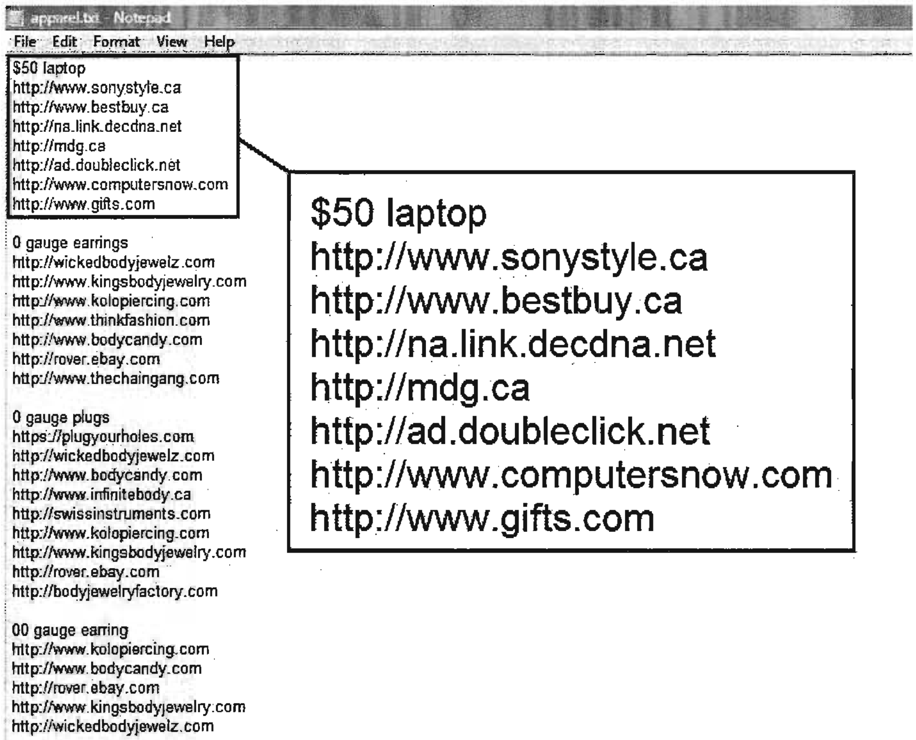

Figure 4.9: Sample output for Python code. After each keyword there is a list of the URL addresses, these URL are ordered according to their appearance in the result page of inquiry for that keyword.

- Gets the URL addresses that appeared after this keyword, and rank them top-down;
- Makes a vertex for each of these URLs if no vertex has already been created for this URL;
- Connect these vertices according to their ranks. Connects a vertex with rank i to all the vertices of ranks smaller than i, if they were not connected to each other before.

This code has two outputs, one a list of vertices and one a list of edges. Each line of the list of vertices contains a number and a URL address. For example, the first line of Figure 4.10 shows that the name of vertex 1 is "http://sonystyle.ca". Each line of the list of edges contains two numbers. The first number is the "tail" of the edge, and the second number is the "head" of that edge. For example, the first line of the Figure 4.11 shows the existence of an edge from vertex 2 ("http://bestbuy.ca") to vertex 1 ("http://sonystyle.ca"). The reason that we made the outputs of C++ code in this format is that we will be using *Pajek* to analyze our graph and the input format of this program is as shown in Figure 4.12.

Figure 4.10: A sample of vertex output of C++ code. Each line consists of the number and name of a vertex.

Pajek is a software for analyzing large-scale networks. Google advertisements network contains 81,791 vertices and 2,112,204 edges (all the data are for August 2010). You can see some pictures of this graph for category "Finance" in Figures 4.13 and 4.14. Because of the high number of vertices in the graph for all keywords, we are not able to draw it with Pajek.

To verify that this network is a small-world network, we need to find the average distance of the graph among reachable pairs of vertices, and the clustering coefficient of the graph. As we said before a small-world network has a small average distance and a high clustering coefficient.

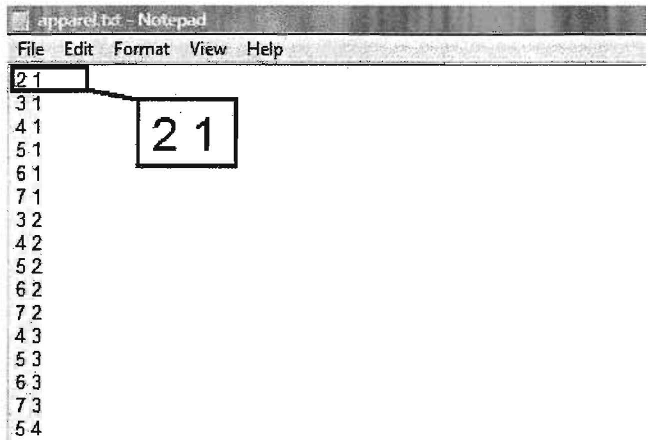

Figure 4.11: A sample of edge output of C++ code. The first line shows the existence of an edge "from" vertex number 2 "to" vertex number 1.

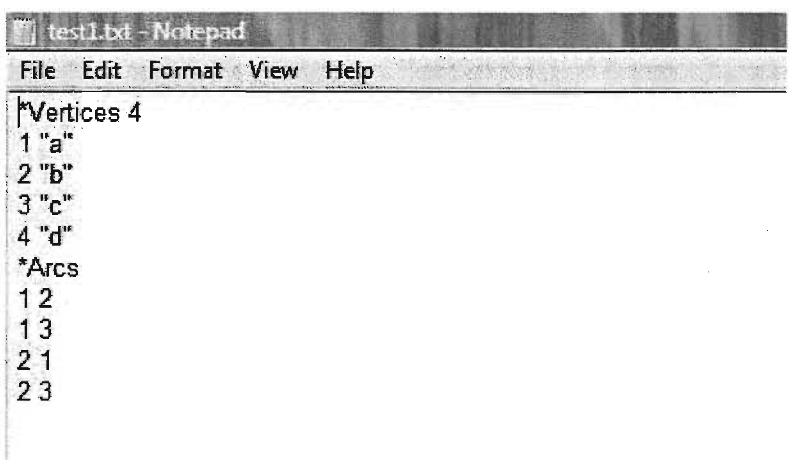

Figure 4.12: A sample input of Pajek program. It shows a directed graph with 4 vertices a,b,c,d and 4 edges.

	category	#of keywords	#of vertices	#of edges	average distance	maximum distance
1	Apparel	48,011	14,255	259,945	2.81201	9
2	Beauty	35,003	9,346	158,964	2.89468	10
3	Computer	50,564	16,705	305,824	3.11624	10
4	Consumer	31,638	7,894	115,783	3.06786	10
5	Family	3,867	1,961	28,905	2.61010	7
6	Finance	37,838	7,965	128,593	3.04893	9
7	Food	38,989	8,352	115,849	2.88312	8
8	Gift	22,203	4,707	80,312	2.74873	8
9	Health	63,995	14,511	245,836	3.03537	9
10	Hobbies	33,775	12,697	173,184	2.86149	10
11	Home	67,921	20,459	455,767	2.78072	10
12	Law	4,189	1,710	15,933	3.02826	8
13	Media	9,069	6,843	59,053	3.30023	12
14	Real state	18,260	7,546	106,076	3.03802	10
15	sport	38,022	10,150	142,747	2.92600	9
16	Travel	22,187	11,000	179,283	2.95928	12
17	Vehicle	36,315	10,129	160,609	2.83656	10
	all categories	561,846	81,791	2,112,204	2.88431	11

Table 4.2: Average distance of Google advertisements network for all categories of keywords.

In the Google advertisements network, %18.2 of the pairs of the vertices are unreachable. The average distance among reachable pairs is 2.88431. The maximum distant among vertices is 11. There are three pairs of vertices of distance 11. One of the pairs consists of vertices "http://GoHomePro.com" (a home inspection) and "http://starschoice.ca" (a limousine service company). We observe that this network has a small average distance. The distance distribution of this graph is shown in Figure 4.15. You can also see the average distance for each category of keywords (17 categories) in Table 4.2.

This network has clustering coefficient of 0.5457577, which is relatively very high. The clustering coefficient distribution is shown in Figure 4.16. The clustering coefficients for each category are shown in Table 4.3.

According to these results, Google advertisements network is a small-world network. Now we look at the degree distribution of this network (we consider the

	category name	clustering coefficient
1	Apparel	0.5561137
2	Beauty	0.5513774
3	Computer	0.4926550
4	Consumer	0.5394617
5	Family	0.5940629
6	Finance	0.5222928
7	Food	0.5353787
8	Gift	0.5662280
9	Health	0.5142095
10	Hobbies	0.5545975
11	Home	0.5675226
12	Law	0.4857277
13	Media	0.5005706
14	Real state	0.5225719
15	sport	0.5598814
16	Travel	0.5808564
17	Vehicle	0.5630946
	all categories	0.5457577

Table 4.3: Clustering coefficient of Google advertisements network for all categories of keywords

in-degree distribution). In-degree distribution of this network has power-law with degree exponent of $\gamma \sim 1.5$, therefore it is a scale-free network (Figure 4.17 plots the in-degree distribution of this network). The maximum degree in this network is 22,574 which belongs to "http://rover.ebay.com" (see Table 4.4).

	vertex name	degree
1	http://rover.ebay.com	22,574
2	http://go.sp-ask.com	17,042
3	http://pronto.com	11,961
4	http://nextag.com	11,140
5	http://na.link.decdna.net	10,609
6	http://info.com	9,972
7	http://best-price.com	9,503
8	http://sears.ca	8,122
9	http://clickserve.dartsearch.net	6,665
10	http://clk.atdmt.com	6,319

Table 4.4: Top 10 vertices with highest degrees in Google advertisements network.

4.2.1 Further characteristics

In the previous section, we formed the Google advertisements graph by considering both top-of-page and sidebar advertisements. Here we make the *Google sidebar advertisements* network by just considering the advertisements who appear on the sidebar of the result page of inquiries. The process of making this graph is the same as before, except that we number the first slot in the sidebar 1, and the second one 2, etc. Clearly, we will have less vertices than before. This graph contains 74,530 vertices and 1,532,130 edges.

The percentage of unreachable pairs is 22.29, and the average distance among reachable pairs is 2.99396 (see Figure 4.19 for distance distribution). We can see an increase in the average of distance. This is because "hubs" are mostly those advertisers who appear on the top-of-page slots. One of the most distant pairs of vertices are "http://moljewelry.com" and "http://view-box.com" with distance of 10 (there are 27 pairs of vertices of distance 10). The clustering coefficient is 0.5105731 (see Figure 4.20 for clustering coefficient distribution). The highest degree is 14,246 which belongs to "http://rover.ebay.com" again (see Table 4.5). In-degree distribution of this network has the power-law with degree exponent of $\gamma \sim 1.6$ (see Figure 4.21).

	vertex name	degree
1	http://rover.ebay.com	14,246
2	http://go.sp-ask.com	13,980
3	http://pronto.com	9,682
4	http://nextag.com	9,024
5	http://na.link.decdna.net	8,958
6	http://info.com	7,409
7	http://best-price.com	6,519
8	http://altfarm.mediaplex.com	5,144
9	http://clickserve.dartsearch.net	4,800
10	http://canadiantire.ca	4,675

Table 4.5: Top 10 vertices with highest degrees in Google sidebar advertisements network.

If we consider only top-of-page advertisements, we will have the *Google top-of-*

page advertisements network. This network shows the following properties: There are 42,888 vertices and 197,552 edges. You can see the number of vertices dropped even more than Google sidebar advertisements graph as Google shows more advertisers in sidebar of the page rather than top of the page. The percentage of unreachable pairs is 57.09, and the average distance among reachable pairs is 3.97289 (see Figure 4.22). the maximum distance is 14 and there are 14 such pairs of vertices such as "http://hoopskills.com" and "http://faro.com". There are 14 pairs of distance 14. The clustering coefficient is 0.3629503 (see Figure 4.23). A list of 10 vertices of highest degree is listed in Table 4.6. The maximum degree is 3,868 and belongs to "http://rover.ebay.com" as in our other graphs. In-degree distribution has the power-law with degree exponent of $\gamma \sim 1.9$ (see Figure 4.24).

	vertex name	degree
1	http://rover.ebay.com	3,868
2	http://clickserve.dartsearch.net	2,141
3	http://na.link.decdna.net	2,105
4	http://sears.ca	1,246
5	http://pronto.com	1,121
6	http://clk.atdmt.com	1,102
7	http://homedepot.ca	1,056
8	http://tracking.searchmarketing.com	758
9	http://track.searchignite.com	741
10	http://go.sp-ask.com	700

Table 4.6: Top 10 vertices with highest degrees in Google top-of-page advertisements network.

We can make the undirected graph of the Google top-of-page advertisements network by making each edge bidirection. The number of vertices is the same as directed graph for this network and the number of edges is 171,863. The percentage of the unreachable pairs is 24.6, and the average distance is 3.58906 (see Figures 4.25). The most distant vertices are "http://fontlab.com" and "http://thelittleappfactory.com" with distance of 14.. The clustering coefficient is 0.5334944 (see Figure 4.26). The highest degree is 7,382. The degree distribution does not have a "typical" power-law property (see Figure 4.27).

Recall that in the third step of the algorithm in making the graph of Google advertisements network, we connect a vertex with rank i to all the vertices of ranks smaller than i, if they were not connected to each other before. If we modify this by adding an edge regardless of existence of an edge, we form a graph with multiple edges. By allowing multiple edges, in the undirected graph of Google top-of-page advertisements network, we obtain a structurally different graph with the following properties. Clearly the number of vertices and the distance of vertices do not change. The number of edges increases to 499,894. The highest degree is 40,262 and as one can see in Figure 4.28, the degree distribution does not have the power-law. The clustering coefficient is not defined for graphs with multiple edges.

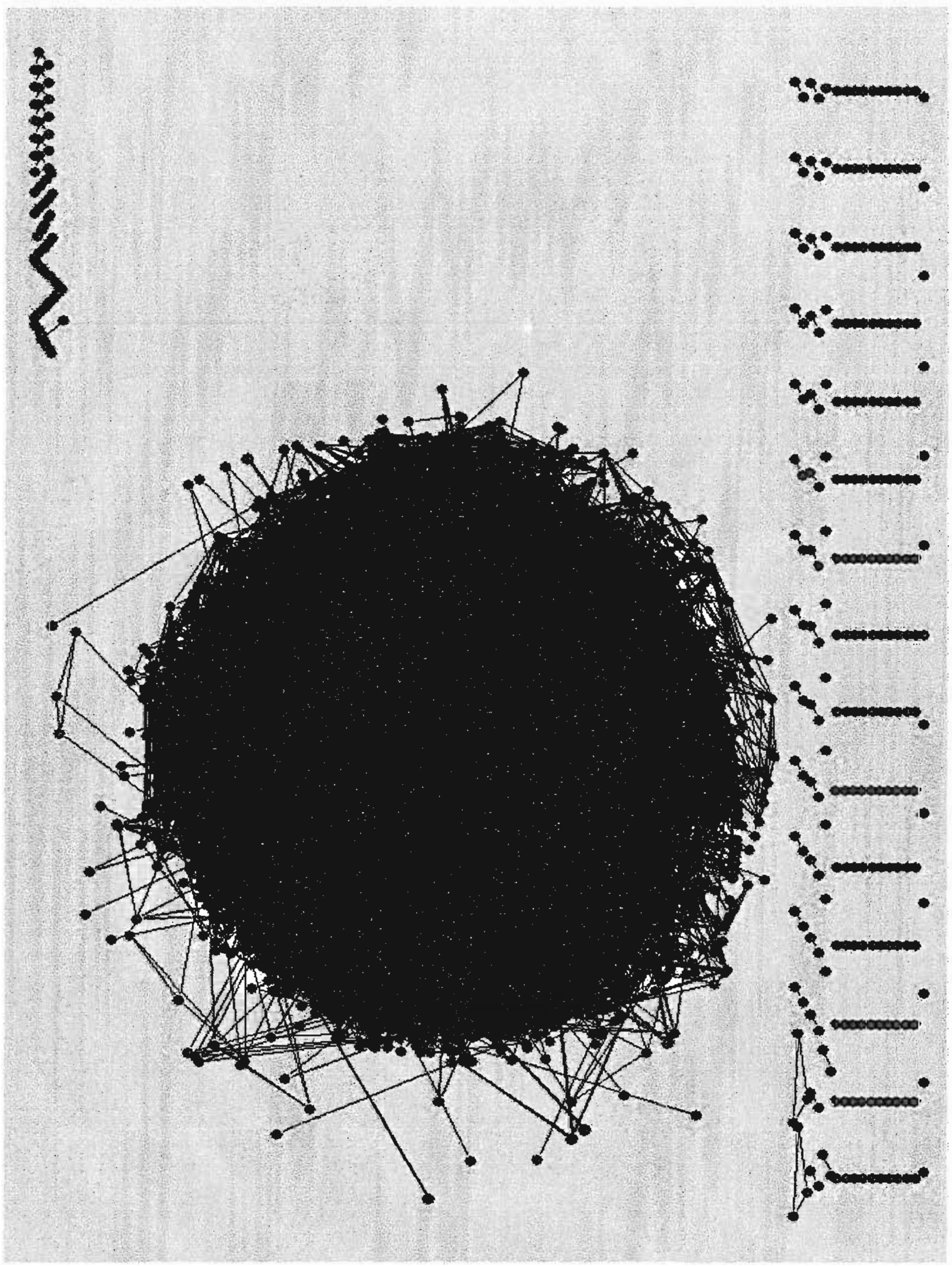

Figure 4.13: Graph of Google advertisements network for "Finance" category. It consists of a giant component and lots of vertices of degree at most 2 which are not connected to this component.

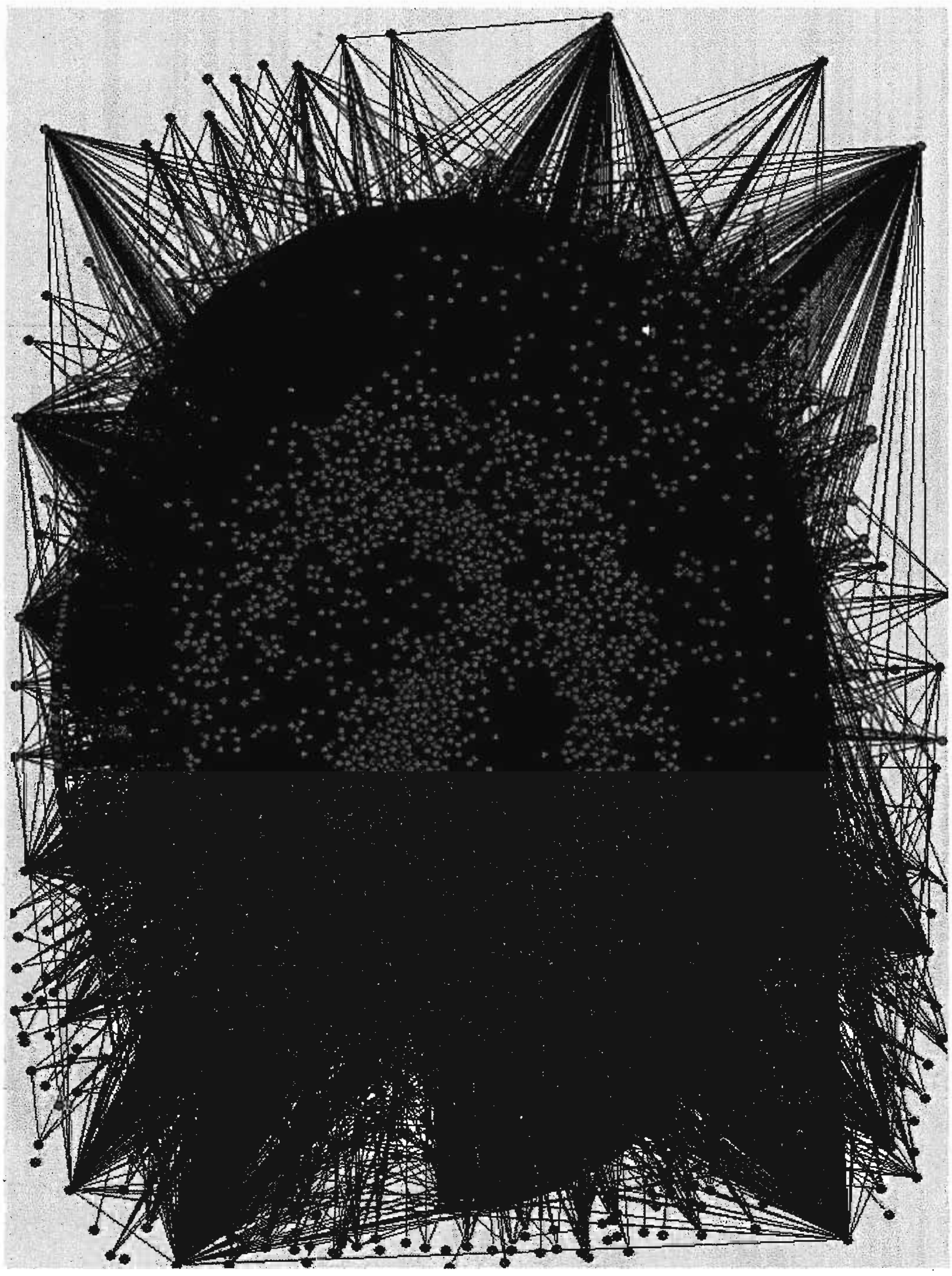

Figure 4.14: Graph of giant component of Google advertisements network for "Finance" category.

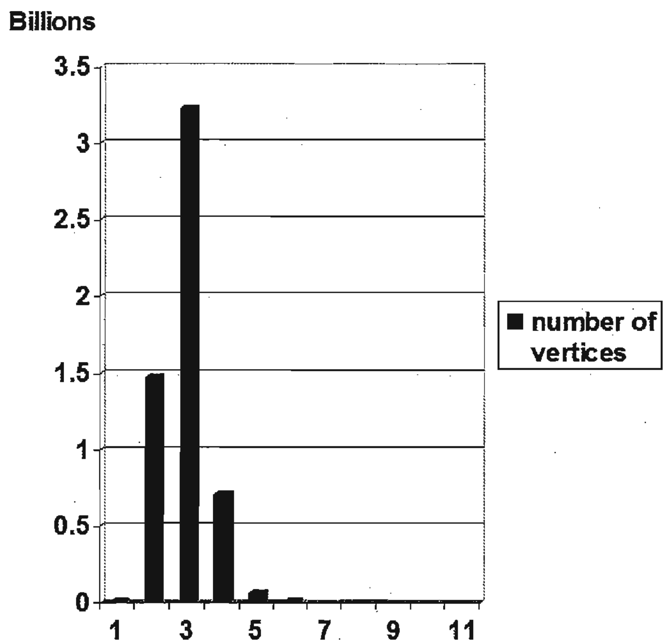

Figure 4.15: Distance distribution of Google advertisements network. X-axis is the distance between two vertices and Y-axis is the number of the pairs with that distance.

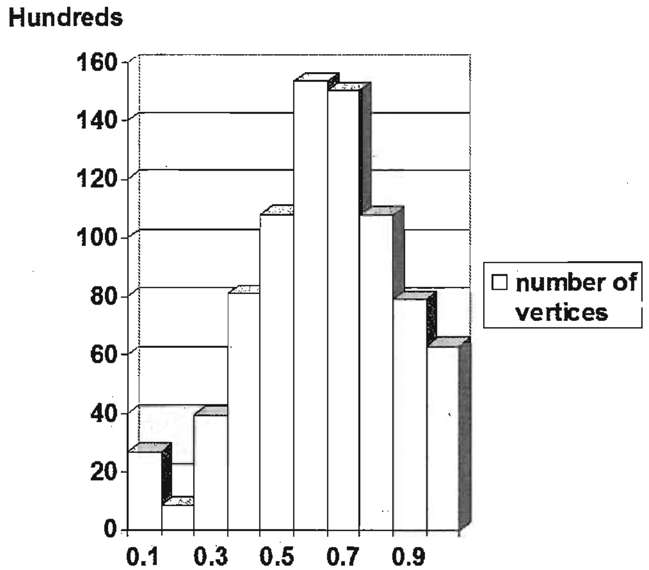

Figure 4.16: Distribution of clustering coefficient of Google advertisements network. X-axis is the clustering coefficient and Y-axis is the number of vertices with that clustering coefficient.

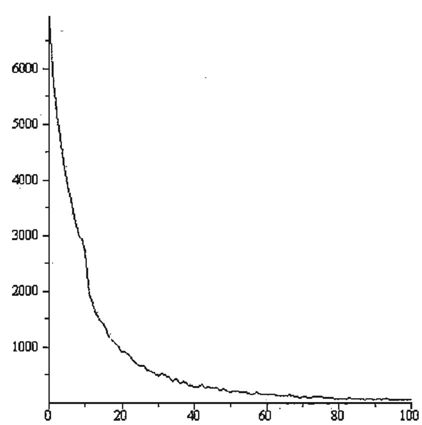

Figure 4.17: In-degree distribution of Google advertisements network. X-axis is the degree and Y-axis is the number of vertices with that degree.

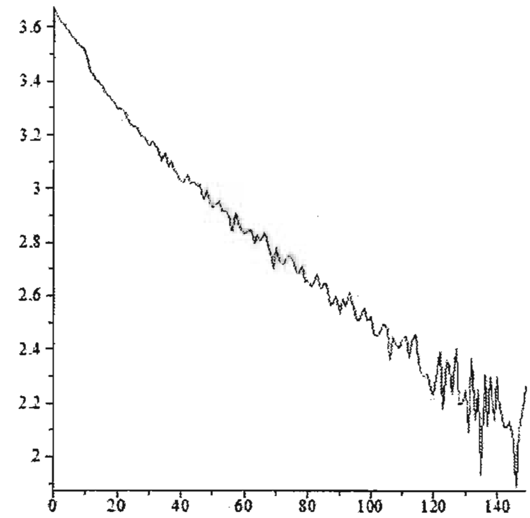

Figure 4.18: LogLog In-degree distribution of Google advertisements network.

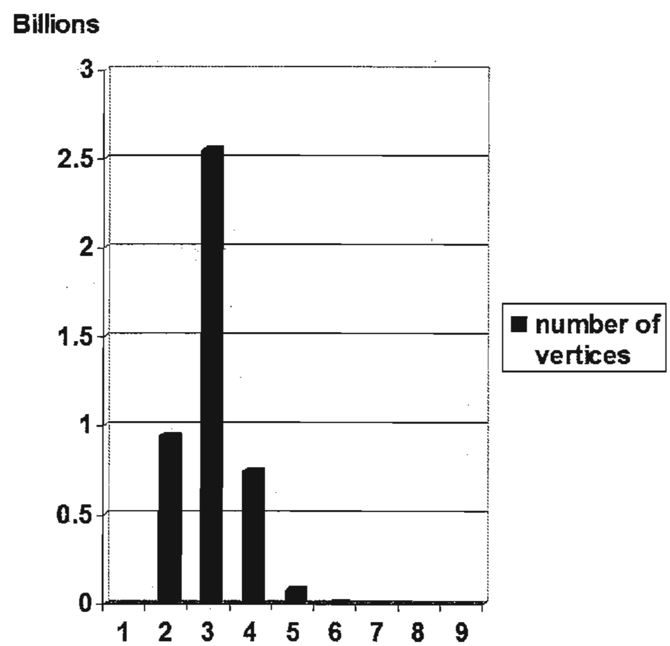

Figure 4.19: Distance distribution of Google sidebar advertisements network. X-axis is the distance between two vertices and Y-axis is the number of the pairs with that distance.

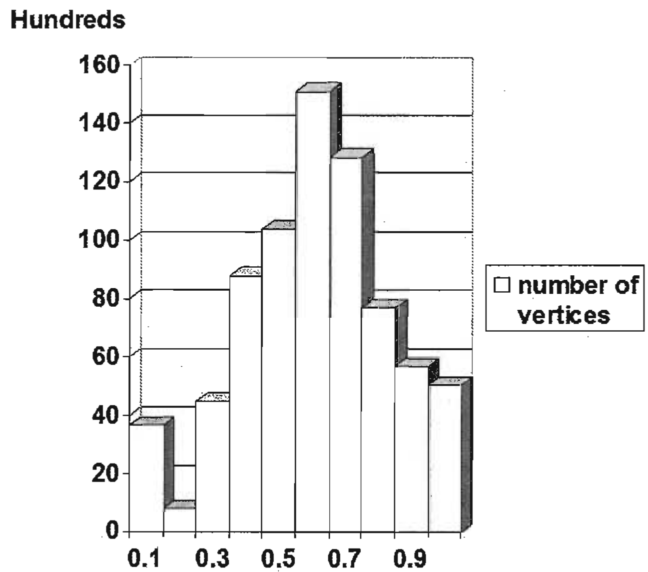

Figure 4.20: Distribution of clustering coefficient of Google sidebar advertisements network. X-axis is the clustering coefficient and Y-axis is the number of vertices with that clustering coefficient.

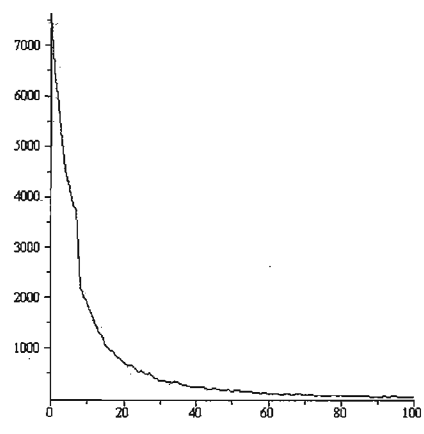

Figure 4.21: In-degree distribution of Google sidebar advertisements network. X-axis is the degree and Y-axis is the number of vertices with that degree.

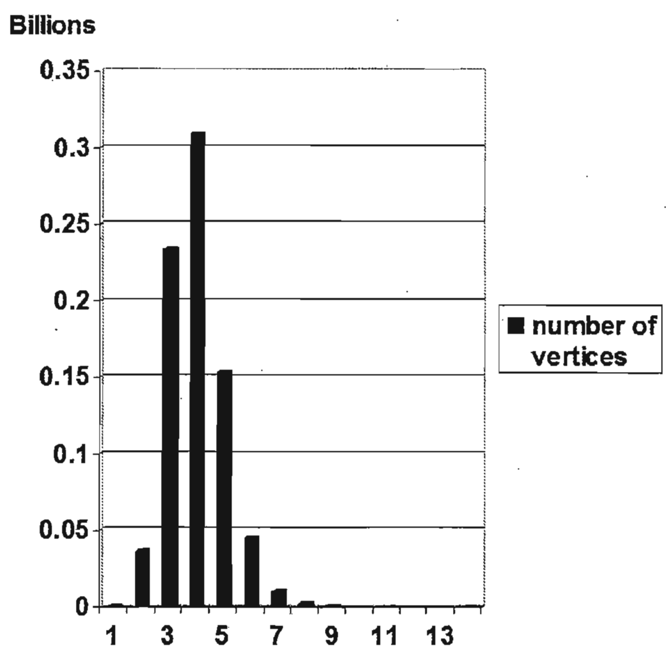

Figure 4.22: Distance distribution of Google top-of-page advertisements network. X-axis is the distance between two vertices and Y-axis is the number of the pairs with that distance.

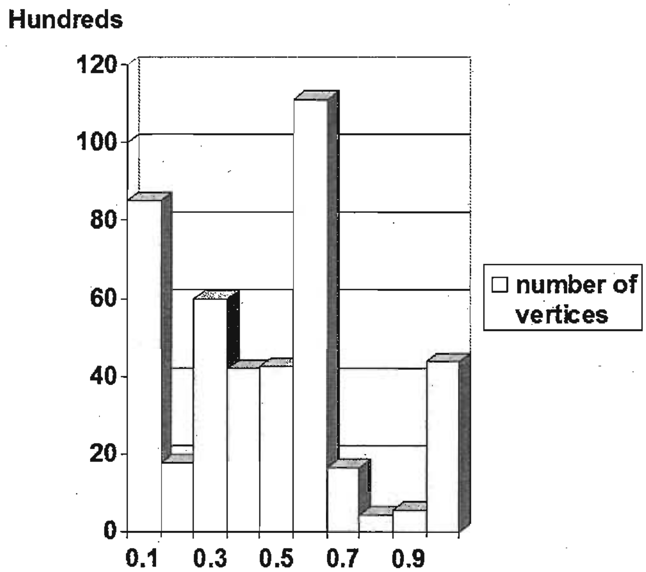

Figure 4.23: Distribution of clustering coefficient of Google top-of-page advertisements network. X-axis is the clustering coefficient and Y-axis is the number of vertices with that clustering coefficient.

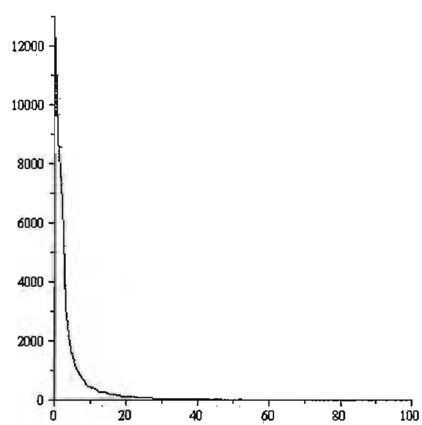

Figure 4.24: In-degree distribution of Google top-of-page advertisements network. X-axis is the degree and Y-axis is the number of vertices with that degree.

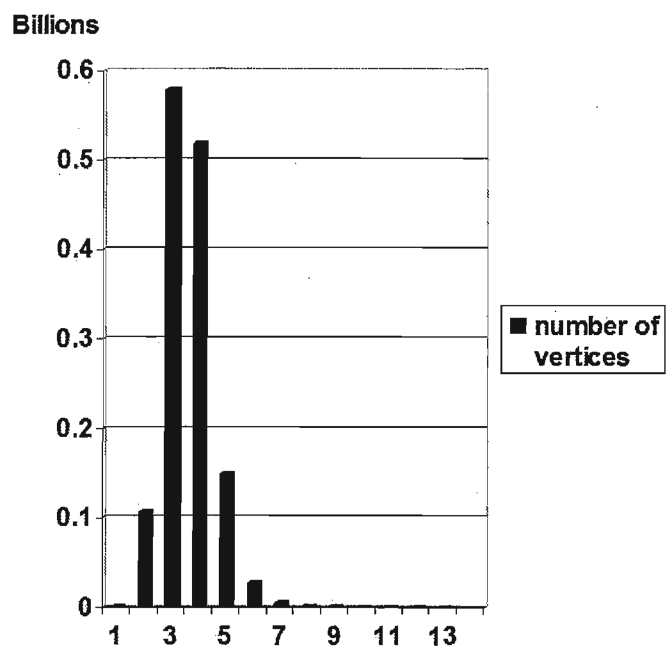

Figure 4.25: Distance distribution of undirected graph of Google top-of-page advertisements network. X-axis is the distance between two vertices and Y-axis is the number of the pairs with that distance.

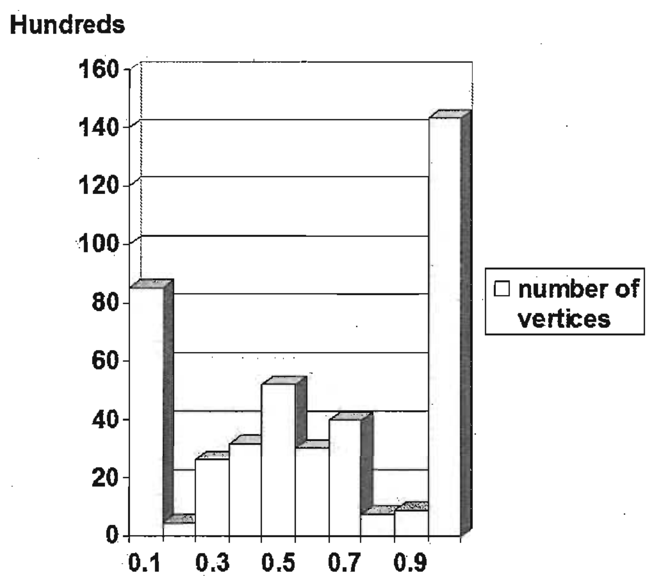

Figure 4.26: Distribution of clustering coefficient of the undirected graph of Google top-of-page advertisements network. X-axis is the clustering coefficient and Y-axis is the number of vertices with that clustering coefficient.

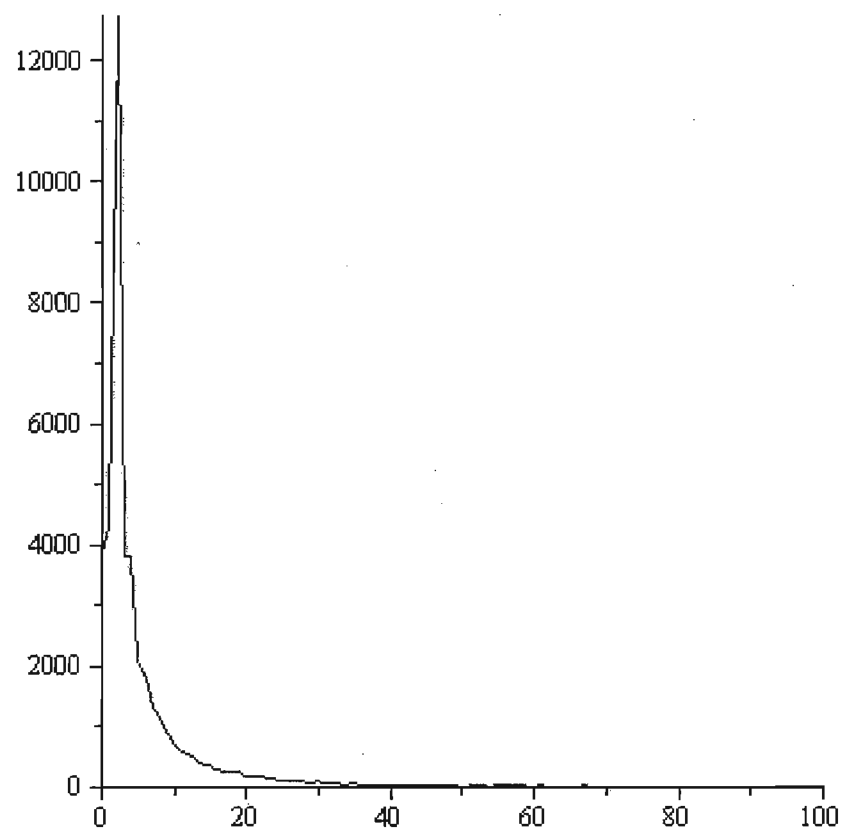

Figure 4.27: Degree distribution of undirected graph of Google top-of-page advertisements network. X-axis is the degree and Y-axis is the number of vertices with that degree.

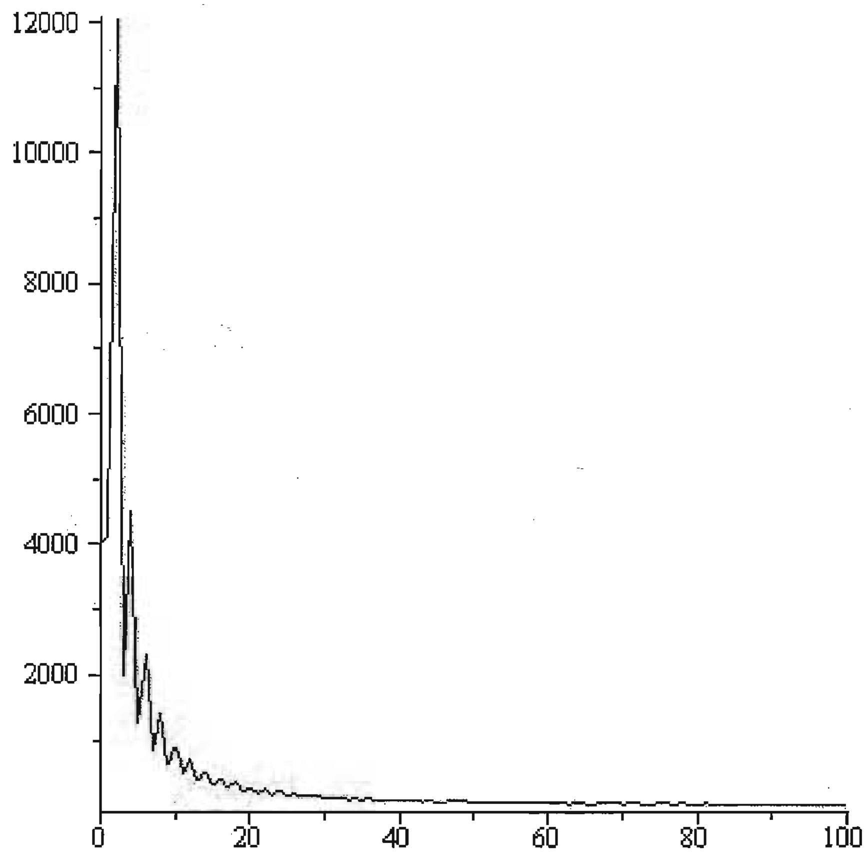

Figure 4.28: Degree distribution of Google advertisements network. Multiple edge undirected graph for up advertisements. X-axis is the degree and Y-axis is the number of vertices with that degree.

Chapter 5

Conclusion

In the theoretical part of this thesis we presented an elementary proof of the celebrated von Neumann minimax theorem for quasiconcave/convex and lower/upper semicontinuous functions, based on the separation of disjoint point and convex set in Euclidean spaces. The minimax theorem being historically connected to fundamental theorems in nonlinear analysis such as the KKM principle, Brouwer fixed point theorem, fixed point and coincidence for set-valued maps, and systems of nonlinear inequality, we undertook a study of implications between these results. Some of these implications are less well-known than others.

We were hoping to provide an elementary proof of the KKM principle. It turns out that such a proof is yet to be written, but our investigations led us to clarify certain aspects of the relationship between intersection theorems of Klee type and KKM type. It turns out that, while the equivalence of the Berge-Klee theorem and the KKM principle is yet to be established, we can say that it is equivalent to a particular version of the KKM principle. This version which we call the convex KKM Principle, holds in arbitrary topological vector spaces thus improving a theorem of Granas and Lassonde. What is known at this moment, is that the equivalence between KKM and Berge-Klee holds if one adopts an approach based on deeper topological tools (with connectedness instead convexity) as in Horvath-Lassonde.

In the experimental part, we determine that considering a basic externality in a model of on-line auctions affects the property of the model. Other kinds of externality can be considered in a future study, for instance:

- Location-dependent externality; such as assuming that a random user only looks at the advertisements in the top X slots, where X is a random variable with a given distribution.
- Long-term externality; if a user finds the advertisements displayed on the result page of her inquiry helpful, she is more likely to click on advertisements in the future. Conversely if the advertisements are found to be not relevant, the user will pay less attention to advertisements in future. This externality sometimes referred to as "the rotten-apple theory of advertising".

We believe that the choice of the externality is a major issue in the study of a sponsored search. So far, this issue has not received enough attention from the research community.

We have also analyzed the network of Google advertisements in terms of average distance, clustering coefficient and degree distribution. We reported that, when this network is modeled as a directed graph, it is small-world scale-free. We showed that the degree distribution of the undirected graph of this network does not have a "typical" power-law property. It would be interesting to find a natural random process that generates a graph with the degree distribution observed in the undirected case.

We made this network several times, based on the data that we gathered in different time intervals. Our results were very similar in all the cases. It is interesting to study the dynamics of the graph, or in other words, how the graph changes over the course of the time. Also further research could compare the Google advertisements graph with other graphs modeling product competition among providers (if they exist), and verifying if the online behavior of advertisers are any different than that of their traditional marketings.

Bibliography

[1] L. A. Adamic. The small world web. *proc. ECDL* 1696 (1999) 443-452. [cited at p. 52]

[2] Y. Y. Ahn, S. Han, H. Kwak, S. Moon, and H. Jeong. Analysis of topological characteristics of huge online social networking services. *Proc. WWW* (2007) 835-844. [cited at p. 52, 53]

[3] R. Albert, H. Jeong, and A. L. Barabási. The diameter of the world wide web. *Nature* 401 (1999) 130-131. [cited at p. 52]

[4] C. D. Aliprantis and K. C. Border. *Infinite Dimensional Analysis.* A Hitchhiker's Guide, 3rd ed., Springer, 2007. [cited at p. 17]

[5] L. A. N. Amaral, A. Scala, M. Barthélémy, and H. E. Stanley. Classes of small-world networks. *PNAS* 97 (2000) 11149-11152. [cited at p. 48]

[6] J. P. Aubin and H. Frankowska. *Set-valued Analysis.* Birkhauser, 1990. [cited at p. 1]

[7] A. L. Barabási and R. Albert. Emergence of scaling in random networks. *Science* 286 (1999) 509512. [cited at p. 54]

[8] A. L. Barabási, H. Jeong, Z. Neda, E. Ravasz, A. Schubert, and T. Vicsek. Evolution of the social network of scientific collaborations. available online at: http://arxiv.org/abs/cond-mat/0104162. [cited at p. 51]

[9] H. Ben-El-Mechaiekh, P. Deguire, and A. Granas. Points fixes et coincidences pour les applications multivoques II (Applications φ et φ^*). *C. R. Acad. Sci. Paris* 295 (1982) 381-384. [cited at p. 1, 2, 16, 22, 23, 25, 26, 27, 28]

[10] H. Ben-El-Mechaiekh and R. W. Dimand. A simple proof of the von Neumann minimax theorem. *Amer. Math. Monthly* in print. [cited at p. 1, 2, 5, 8, 19]

[11] H. Ben-El-Mechaiekh and R. W. Dimand. *The von Neumann minimax theorem revisited. Banach Center* 77 (2007) 23-34. [cited at p. 2]

[12] H. Ben-El-Mechaiekh and R. W. Dimand. Von Neumann, Ville, and the minimax theorem. *Int. Game Thy. Rev.* 12 (2010) 115-137. [cited at p. 2, 3, 27]

[13] C. Berge. Sur une propriete combinatoire des ensembles convexes. *C.R. Acad. Sci. Paris* 248 (1959) 2698-2699. [cited at p. 1, 3, 18]

[14] B. Bollobás. Mathematical results on scale-free random graphs. January 16, 2003. [cited at p. 49]

[15] F. E. Browder. The fixed point theory of multi-valued mappings in topological vector spaces. *Math. Ann.* 177 (1968) 283-301. [cited at p. 1, 2, 26]

[16] J. Dugundji and A. Granas. KKM maps and variational inequalities. *Ann. Scuola Norm. Sup. Pisa* 5 (1978) 679-682. [cited at p. 1, 2, 15, 23, 26, 27]

[17] K. Fan. A generalization of Tychonoff's fixed point theorem. *Math. Ann.* 142 (1961) 305310. [cited at p. 16, 22]

[18] K. Fan. Applications of a theorem concerning sets with convex sections. *Math. Ann.* 163 (1966) 189-203. [cited at p. 3]

[19] K. Fan. some properties of convex sets related to fixed point theorems. *Math. Ann.* 266 (1984) 519-537. [cited at p. 22]

[20] A. Ghouila-Houri. Sur l'étude combinatoire des families de convexes. *C. R. Acad. Sci. Paris* 252 (1961), 494-496. [cited at p. 32, 33]

[21] R. Gomes, N. Immorlica, and E. Markakis. Externalities in keyword auctions:an empirical and theoretical assesment. *Proc. WINE* (2009) 172183. [cited at p. 38]

[22] A. Granas. Travaux mathématiques. Partie I: Analyse Fonctionelle. [cited at p. 30]

[23] A. Granas and M. Lassonde. Some elementary general principles of convex analysis. *Top. Meth. Nonlinear. Anal.* 5 (1995) 23-37. [cited at p. 2, 8, 32]

[24] C. D. Horvath and M. Lassonde. Intersection of sets with n-connected unions. *Proc. Amer. Math. Soc.* 125 (1997) 1209-1214. [cited at p. 2, 16, 31, 32]

[25] I. Joó. Simple proof for von Neumann's minimax theorem. *Acta Sci. Math.* 42 (1980) 91-94. [cited at p. 2]

[26] J. Kindler. A simple proof of Sion's minimax theorem. *Amer. Math. Monthly* 112 (2005) 356358. [cited at p. 2]

[27] T. H. Kjeldsen. John von Neumann's conception of the minimax theorem: A journey through different mathematical contexts. *Archive for History of Exact Sciences* 56 (2001) 39-68. [cited at p. 2, 14]

[28] V. L. Klee. On certain intersection properties of convex sets. *Canad. J. Math.* 3(1951) 272-275. [cited at p. 1, 18]

[29] B. Knaster, K. Kuratowski, and S. Mazurkiewicz. Ein beweis des fixpunktsatzes fur n-dimensionale simplexe. *Fund. Math.* 14 (1929) 132137. [cited at p. 1, 15]

[30] F. Liljeros, C. R. Edling, L. A. N. Amaral, H. E. Stanley, and Y. Aberg. The web of human sexual contacts. *Nature* 411 (2001) 907-908. [cited at p. 52]

[31] M. Magill and M. Quinzii. *Theory of Incomplete Markets.* vol. 1, The MIT Press, 1996. [cited at p. 10, 11]

[32] S. Milgram. The small world problem. *Psychology Today* 1 (1967) 61-67. [cited at p. 48]

[33] M. E. J. Newman. Scientific collaboration networks. i. network construction and fundamental results. *Phys. Rev. E.* 64 (2001) 016131-1 - 016131-8. [cited at p. 51]

[34] M. E. J. Newman. The structure of scientific collaboration networks. *Proc. Natl. Acad. Sci.* 98 (2001) 404-409. [cited at p. 51]

[35] H. Nikaido. On von Neumann's minimax theorem. *Pacific J. Math.* 4 (1954) 65-72. [cited at p. 1, 4, 21]

[36] N. Nisan, T. Roughgarden, E. Tardos, and V. Vazirani. *Algorithmic Game Theory.* Cambridge University Press, New York, 2007. [cited at p. 37]

[37] S. Park. The Knaster-Kuratowski-Mazurkiewicz theorem and almost fixed points. *T. M. N. A.* 16 (2000) 195-200. [cited at p. 32]

[38] H. Schaefer, M. Wolff, and M. P. H. Wolff. *Topological Vector Spaces.* 2nd ed. Springer Verlag, New York, 1999. [cited at p. 13]

[39] M. Sion. On general minimax theorems. *Pacific J. Math.* 8 (1958) 171-176. [cited at p. 1, 4, 21]

[40] H. R. Varian. Position auctions. *Int. J. Ind. Organ.* 25 (2007) 1163-1178. [cited at p. 40]

[41] J. von Neumann. Sur la théorie des jeux. *C. R. Acad. Sci. Paris* 186 (1928) 1689-1691. [cited at p. 3]

[42] J. von Neumann. Uber ein okonomisches gleichungssystem und eine verallgemeinerung des Brouwerschen fixpunktsatzes. *Ergeb. eines Math. Koll. Vienna (ed. by Karl Menger)* 8 (1937) 73-83. [cited at p. 4]

[43] J. von Neumann. Zur theorie der gesellschaftsspiele. *Math. Ann.* 100 (1928) 295-320. [cited at p. 4]

[44] S. Wasserman and K. Faust. *Social Network Analysis: Methods and Applications.* Cambridge University Press, New York, 1994. [cited at p. 48]

[45] D. B. West. *Introduction to Graph Theory.* Prentice Hall, Inc., Upper Saddle River, NJ, 1996. [cited at p. 45]

Appendices

Appendix A

Python code

```
def myfunc():
 import os
 import httplib
 ff=open('C:\\Users\\sama\\Desktop\\ListOfKeywords.txt','r')
 conn=httplib.HTTPConnection("www.google.com")
 words=ff.readlines()
 for g in words:
  print g
  conn.request("GET","/search?q="+g.strip().replace(' ','%20'))
  ss=conn.getresponse()
  kk=ss.read()
  gg=open('C:\\Users\\sama\\Desktop\\PythonResult.txt','a')
  gg.write(g)
  #Beginning of the sponsored link
  pos1=kk.find('a id=pa')
  if pos1<0:
   print "There is no sponsored links"
```

```
    #Here in the textfile we need to print the keyword and
    #go to 2nd next line for the next keyword
    gg.write('\n')
  else:
    myFlag = 0
    while myFlag == 0:
      httpPos1=kk.find('http',pos1)
      httpPos2=kk.find('/',httpPos1+1)
      httpPos3=kk.find('/',httpPos2+1)
      httpPos4=kk.find('/',httpPos3+1)
      httpPos5=kk.find('"',httpPos3+1)
      httpPos6=kk.find('%',httpPos3+1)
      if httpPos6 != -1 and httpPos6 < httpPos5:
        httpPos5=httpPos6
      if httpPos5 < httpPos4:
        httpPos4=httpPos5
      myStr=kk[httpPos1:httpPos4]
      gg.write(myStr + '\n')
      pos1=kk.find('a id=pa',httpPos4)
      if pos1<0:
        myFlag = 1
    gg.write('\n')
  gg.close()
```

Appendix B

C++ code

```
#include "stdafx.h"
#include <iostream>
#include <fstream>
#include <string>
using namespace std;
class Vertex;
class Edge;
int countnode=0; // number of total vertices of the graph
int nodenumber=1; // number of each vertex

class Vertex {
 public:
  Vertex(string, int, Vertex *);
  ~Vertex();
  string getData();
  int getNumber();
  Vertex *getNext();
```

```
  Edge *getFirstEdge();
  void connectTo(Vertex *);
  bool HaveEdgeTo(Vertex*);
 private:
  string data;//stores name of the advertisers
  int Number;//stores number of the vertex
  Edge *edges;//pointer to the list of out coming edges of the vertex
  Vertex *next;//pointer to the next vertex in the graph
};
class Edge{
 public:
  Edge(Vertex *, Edge *);
  ~Edge();
  Vertex *getEnd();
  Edge *getNext();
 private:
  Vertex *end; // pointer to the vertex related to this edge
  Edge *next; // pointer to the next edge in the list of edges
              // associated with the vertex
};
class Graph {
 public:
  Graph();
  ~Graph();
  Vertex *AddVertex(string);
  bool HavingEdge(Vertex*,Vertex*);
  void AddEdge(Vertex*,Vertex*);
  int countingnodes();
```

```
  Vertex *findVertex(string);
  int printVertexNumber();
  string printvertices();
 private:
  Vertex *first; // a pointer to the first vertex of the graph
};

// ************** methods of class Vertex *********************

// Constructor
Vertex::Vertex(string theData, int theNumber, Vertex *nextVertex){
  data = theData;
  Number = theNumber;
  next = nextVertex;
  edges = NULL;
}
// Destructor
Vertex:: ~Vertex() {
  delete next;
  delete edges;
}
// returns the data of the vertex
string Vertex::getData() {
  return data;
}
// return the number of the vertex
int Vertex::getNumber() {
  return Number;
```

```
}
// returns the pointer to the next vertex in the graph
Vertex *Vertex::getNext() {
  return next;
}
// returns the pointer to the first out coming edge of the vertex
Edge *Vertex::getFirstEdge() {
  return edges;
}
//adds an edge to connect the vertex to the vertex pointed to by A
void Vertex::connectTo(Vertex *A) {
  // allocate memory for a new Edge,set its Vertex pointer to point
  // to A, and its Edge pointer to point to the rest of edges
  Edge *newEdge = new Edge(A, edges);
  edges = newEdge;// make the new edge the first edge of the vertex
}
//checks if there is an edge between vertex and vertex pointed by Q
bool Vertex::HaveEdgeTo(Vertex *Q){
  Edge *W=edges;
  while (W != NULL) {
          if (W->getEnd()->getData() == Q->getData())
                  return true;
          W = W ->getNext();
  }
  return false;
}
//*********************methods of class Edge*********************
```

```
// Constructor: sets the two fields to the two given values
Edge::Edge(Vertex *Vert, Edge *nextEdge) {
  end = Vert;
  next = nextEdge;
}
// Destructor: calls the destructor for the next edge on the list
Edge::~Edge() {
  delete next;
}
// returns the pointer to the end vertex of the edge
Vertex *Edge::getEnd() {
  return end;
}
// returns the pointer to the next edge on the list
Edge *Edge::getNext() {
  return next;
}
//*********************methods of class graph*****************

// Constructor
Graph::Graph() {
  first = NULL;
}
// Destructor
Graph::~Graph() {
  delete first;
}
// returns the pointer to the vertex with data theData if such
```

```
// vertex occurs in the graph, otherwise returns NULL
Vertex *Graph::findVertex(string theData) {
  Vertex *A = first;
  while (A != NULL) {
    if ( A -> getData() == theData)
      return A;
    A = A -> getNext(); // go to the next vertex
  }
  return NULL;
}
//if there is an edge between vertices A and B returns true,
//otherwise false
bool Graph::HavingEdge(Vertex *A, Vertex *B){
        bool C = A->HaveEdgeTo(B);
        return C;
}
// add a new vertex to the graph with data theData and returns
// the pointer to this new vertex
Vertex *Graph::AddVertex(string theData) {
  // allocate memory for new vertex with data theData,
  //make it point to the previous first vertex
  Vertex *newVertex = new Vertex(theData, nodenumber, first);
  // make the new vertex the first one in the list of vertexes
  first = newVertex;
  nodenumber++;// increase the number of vertices of the graph
  return newVertex;// return the pointer to the new vertex
}
// creates an edge from vertex A to vertex B
```

```
void Graph::AddEdge(Vertex *A, Vertex *B) {
  A -> connectTo(B); // connect A  to B
}
// count the number of vertices of the graph
int Graph::countingnodes() {
  Vertex *A=first; // makes A point to the first vertex of the graph
  while (A != NULL){// while it is not the end of the vertex list do
          countnode++;//increase the number of vertices by 1
          A = A->getNext(); // go to the next vertex
  }
  return countnode;
}
// returns the number of each vertex
int Graph::printVertexNumber(){
        int out = first->getNumber();
        return out;
}
// return the name of the advertisers
string Graph::printvertices(){
        string out=first->getData();
        first=first->getNext();
        return out;
}
//************************************************************
int main() {

int i=0;
int node=0;
```

```
bool yx;
Vertex *x,*y;
string str="www";
string line = "";
string temp[30];
Graph myGraph;
//opens the list of keywords and the name of
// advertisers for each of them
ifstream myfile ("input.txt");
//opens a new file to put the list of Vertices in Pajek format on it
ofstream Vertexfile ("Vertex output.txt");
//opens a new file to put the list of edges in Pajek format on it
ofstream Edgefile ("Edges output.txt");
if (myfile.is_open()){
 if (Edgefile.is_open()){
  // write *Arcs at the beginning of the Edgefile
  Edgefile << "*Arcs " << endl;
  while (! myfile.eof() ){
   getline (myfile, line);
   getline (myfile, line);
   while (!line.empty()){
   // deleting www. from the name of the advertisers
    if  (line.find(str)==7)
     temp[i]=line.erase (7,4);
        else
         temp[i]=line;// save up to 3 advertisers for each keyword
        std::cout << temp[i] << std::endl;
    // if this advertiser doesn't exist in the list of vertices
```

```
    // add it as a new vertex
    x = myGraph.findVertex(temp[i]);
        if (x == NULL)
         myGraph.AddVertex(temp[i]);
        i++;
        getline (myfile, line);
        }//end while (!line.empty())
    // connect lower advertisers to upper one for each keyword
    for (int j=0; j<=i-2;j++){
         x = myGraph.findVertex(temp[j]);
         for (int k=j+1; k<=i-1;k++){
          y = myGraph.findVertex(temp[k]);
          yx = myGraph.HavingEdge(y,x);
          if  (yx != true){
       myGraph.AddEdge(y,x);
       // write edges on Edgefile readable for Pajek
           Edgefile << y->getNumber() << " " << x->getNumber() << endl;
          }//end if
         }//end for
        }//end for
        i=0;
  }// end while (! myfile.eof() )
 }// end if (Edgefile.is_open())
 else cout << "Unable to open output";
}// end if (myfile.is_open())
else cout << "Unable to open input";

node=myGraph.countingnodes();
```

```
if (Vertexfile.is_open()){//creating vertex file readable for Pajek
 Vertexfile << "*Vertices " << node << endl;
 for (int j=1 ; j<=node; j++){
  cout << j << endl;
  Vertexfile << myGraph.printVertexNumber() << " \"";
  Vertexfile << myGraph.printvertices() << "\"" << endl;
 }// end for
}// end if
else cout << "Unable to open Vertexoutput";

int t;
cin>>t;
return 0;
}// end main
```

List of Figures

List of Tables